Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών

Εργαστήριο Λογικής και Επιστήμης Υπολογισμών

Μεγιστοποίηση Κέρδους σε Δημοπρασίες Αναβαλλόμενης Αποδοχής

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ερμής Νικηφόρος Σουμαλιάς

Επιβλέπων: Δημήτριος Φωτάκης
Αναπληρωτής Καθηγητής Ε.Μ.Π.

Αθήνα, 15/09/2020

Εθνικο Μετσοβιο Πολυτεχνειο
Σχολη Ηλεκτρολογων Μηχανικων και Μηχανικων Υπολογιστων
Τομεας Τεχνολογιας Πληροφορικης και Υπολογιστων

Μεγιστοποίηση Κέρδους σε Δημοπρασίες Αναβαλλόμενης Αποδοχής

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ερμής Νικηφόρος Σουμαλιάς

Επιβλέπων: Δημήτριος Φωτάκης
Αναπληρωτής Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 15/09/2020.

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

....
Ερμής Νικηφόρος Σουμαλιάς
(Διπλωματούχος Ηλεκτρολόγος Μηχανικός & Μηχανικός Υπολογιστών Ε.Μ.Π.)

Περίληψη

Στη διπλωματική αυτή, μελετάμε το πρόβλημα της μεγιστοποίησης εισοδήματος σε δημοπρασίες αναβαλλόμενης αποδοχής για την πώληση πολλαπλών αντιτύπων. Από την εισαγωγή τους από τους Milgrom και Segal, οι δημοπρασίες αναβαλλόμενης αποδοχής έχουν χρησιμοποιηθεί εκτενώς χάρη στις αξιοσημείωτες ιδιότητές τους, αναφορικά προφανή φιλαλήθεια και weak group-strategy-proofness. Η περισσότερη έρευνα στις δημοπρασίες αναβαλλόμενης αποδοχής είτε είχε εστιαστεί στο στόχο της μεγιστοποίησης της κοινωνικής ωφέλειας, είτε είχε ακολουθήσει μία προσέγγιση χειρότερης περίπτωσης. Στη διπλωματική αυτή εφαρμόζουμε PAC μάθηση στις δημοπρασίες αναβαλλόμενης αποδοχής και δείχνουμε ότι, υπό φυσικές υποθέσεις για τις κατανομές των παικτών, δοθέντων κάποιων δειγμάτων των κατανομών αυτών είναι εφικτή η μάθηση δημοπρασιών αναβαλλόμενης αποδοχής με υψηλό αναμενόμενο εισόδημα. Εστιάζουμε σε 2 διακεκριμένα περιβάλλοντα: Ένα περιβάλλον μίας παραμέτρου, δημοπρασίες πολλαπλών αντιτύπων με παίκτες με αθροιστικές συναρτήσεις αξίας, και ένα περιβάλλον πολλαπλών παραμέτρων, δημοπρασίες πολλαπλών αντιτύπων με παίκτες με submodular συναρτήσεις αξίας.

Στην περίπτωση των παικτών με αθροιστικές συναρτήσεις αξίας, παρέχουμε μία πρωτότυπη υλοποίηση των δημοπρασιών t-επιπέδων των Morgenstern και Roughgarden ως δημοπρασίες αναβαλλόμενης αποδοχής και, χρησιμοποιώντας την υλοποίηση αυτή, παρέχουμε άνω όρια για την δειγματική πολυπλοκότητα που απαιτείται για τον προσδιορισμό μίας δημοπρασίες αναβαλλόμενης αποδοχής με αναμενόμενο εισόδημα αυθαίρετα κοντά στο βέλτιστο. Το όριό μας με την προσέγγιση αυτή είναι εξίσου καλά με το αντίστοιχο όριο της αρχικής εργασίας. Έπειτα επεκτείνουμε το αποτέλεσμά μας αυτό από δημοπρασίες πολλαπλών αντιτύπων σε δημοπρασίες με αυθαίρετους polymatroid περιορισμούς. Για να το επεκτείνουμε αυτό, επεκτείνουμε τη δομή των δημοπρασιών t-επιπέδων, προσθέτοντας μία επιπλέον παράμετρο ανά επίπεδο, το σκορ του επιπέδου αυτού. Όπως δείξαμε για πολυματροειδή περιβάλλοντα, αυτό ενισχύει σημαντικά τις εκφραστικές ικανότητες της κλάσης ενώ αυξάνει ελάχιστα την ψευδοδιάστασή της. Στην πράξη, αυτό σημαίνει ότι κανείς μπορεί να επιτύχει τον ίδιο λόγο προσέγγισης χρησιμοποιώντας πολύ λιγότερα επίπεδα, και κατά συνέπεια με μικρότερη δειγματική πολυπλοκότητα.

Στην περίπτωση που οι παίκτες έχουν submodular συναρτήσεις αξιών, προτείνουμε την κλάση των unit bundling δημοπρασιών. Η προσέγγισή μας είναι σχετικά πρωτότυπη: Εισάγουμε ένα νέο μηχανισμό, τον "εκ των προτέρων βέλτιστο" μηχανισμό, ο οποίος πάντα επιλέγει την σταθερή ανάθεση με τη μεγαλύτερη αναμενόμενη κοινωνική ωφέλεια. Έπειτα, προκειμένου να φράξουμε προς τα πάνω το σφάλμα αντιπροσώπευσης της unit bundling κλάσης χρησιμοποιούμε την αναμενόμενη κοινωνική ωφέλεια του μηχανισμού αυτού ως μία διεπαφή μεταξύ της αναμενόμενης κοινωνικής ωφέλειας του VCG μηχανισμού και του αναμενόμενου εισοδήματος του βέλτιστου μηχανισμού από την unit bundling κλάση. Συνδυάζοντας αυτά τα δύο φράγματα δείχνουμε ότι, υπό φυσικές υποθέσεις για τις κατανομές των παικτών, δοθέντων κάποιων δειγμάτων των αξιών τους είναι δυνατή η μάθηση μίας δημοπρασίας αναβαλλόμενης απόφασης με αναμενόμενο εισόδημα ίσο με την αναμενόμενη κοινωνική ωφέλεια του VCG, μείον κάποιους μικρούς αθροιστικούς όρους που είναι υπογραμμικοί στον αριθμό των παικτών και τις διασπορές των κατανομών αξιών τους. Το αποτέλεσμα αυτό είναι αξιοσημείωτο δεδομένου ότι στο περιβάλλον αυτό, κανένας weakly group-strategyproof μηχανισμός δεν μπορεί να εγγυηθεί κοινωνική ωφέλεια που είναι πάνω από $\frac{1}{\sqrt{2}}$ προσέγγιση της βέλτιστης.

Τέλος, επεκτείνουμε τα αποτελέσματα μας και για περιβάλλοντα μίας και πολλαπλών μεταβλητών στην περίπτωση όπου ο αριθμός δειγμάτων είναι σημαντικά περιορισμένος. Συγκεκριμένα, για το περιβάλλον μίας παραμέτρου προτείνουμε μία δημοπρασία αναβαλλόμενης απόφασης η οποία, χρησιμοποιώντας ένα μονάχα δείγμα, επιτυγχάνει τουλάχιστον $\frac{1}{4}$ του βέλτιστου αναμενόμενου εισοδήματος. Για την απόδειξη αυτή χρειάστηκε να επεκτείνουμε την έννοια των "ισόμετρων μηχανισμών" που εισήγαγαν οι Hartline και Roughgarden σε αυθαίρετα περιβάλλοντα μίας παρα-

μέτρου. Έπειτα, για το περιβάλλον πολλών παραμέτρων προτείνουμε έναν μηχανισμό ο οποίος, χρησιμοποιώντας 2 δείγματα, έχει αναμενόμενο εισόδημα που είναι τουλάχιστον 0.589 αυτού που είχε ο μηχανισμός στην προσέγγιση δειγματικής πολυπλοκότητας για το ίδιο περιβάλλον. Υπό τις ίδιες φυσικές υποθέσεις για τις κατανομές των παικτών, αυτό είναι 0.589 προσέγγιση της αναμενόμενης κοινωνικής ωφέλειας του VCG, μείον κάποιους μικρούς αθροιστικούς όρους.

Λέξεις κλειδιά: Αυτοματοποιημένος Σχεδιασμός Μηχανισμών, Δημοπρασίες Αναβαλλόμενης Αποδοχής, Δημοπρασίες Πολλαπλών Αντιτύπων, Δημοπρασίες t-επιπέδων, Δειγματική Πολυπλοκότητα.

Abstract

In this thesis, we study the problem of revenue maximization in deferred-acceptance multi-unit auctions. Since their introduction by Milgrom and Segal, deferred-acceptance auctions have been used extensively due to heir remarkable incentive properties, namely weak group-strategyproofness and obvious truthfulness. Most of the work on deferred-acceptance auctions has either focused on the objective of social welfare maximization or adopted a worst-case analysis approach. In this thesis we apply the framework of PAC learning to deferred-acceptance auctions and show that, under natural distributions assumptions, given few samples, it is possible to learn deferred-acceptance auctions with very high expected revenue.

We focus on two distinct environments: a single-parameter environment, multi-unit auctions with bidders with additive valuation functions, and a multi-parameter environment, multi-unit auctions with bidders with submodular valuation functions.

In case of bidders with linear valuations, we provide a novel implementation of Morgenstern and Roughgarden's t-level auctions as deferred-acceptance auctions and, using that implementation, we upper bound the sample complexity of determining a deferred-acceptance auction with expected revenue arbitrarily close to optimal. Our bound using that approach is equally good to that of their original paper.

Then, we extend this result from multi-unit auctions to environments with arbitrary polymatroid constraints. To achieve this, we extend the framework of t-level auctions, adding an additional parameter per level, its level score. As we show for polymatroid environments, this significantly increases the expressive capabilities of the class, while only slightly increasing its pseudo-dimension. In practice, this means that one can achieve the same approximation ratio using fewer levels and thus with decreased sample complexity.

In case of bidders with submodular valuation functions, we propose the class of unit bundling mechanisms with non-anonymous bundle sizes and reserve prices. We introduce a new mechanism, which we call "a priori optimal". The "a priori optimal" mechanism always chooses the fixed allocation with the highest expected social welfare. Then, in order to establish a representation error bound for the unit bundling mechanism class, we use the expected social welfare of this mechanism as an interface between the expected social welfare of VCG and the expected revenue of the optimal mechanism from the unit bundling class. In order to establish a generalization error bound for the mechanism class, we use Balcan's sample complexity framework. Combining these two bounds we prove that, under natural distribution assumptions, given some samples of the bidders' valuations, it is possible to learn a deferred-acceptance auction with expected revenue equal to the expected social welfare of VCG minus some small additive terms that are sublinear in the number of players and the variances of their valuation distributions. We remark that in this environment, no weakly group-strategyproof mechanism can guarantee social welfare that is within a factor of $\frac{1}{\sqrt{2}}$ from the optimal one.

Afterwards, we extend both our results for single- and multi-parameter environments to the case where the number of samples is severely restricted. Specifically, for the single-parameter environment, we propose a mechanism that, using a single sample, achieves on expectation over the draw of that sample an expected revenue within a factor of $\frac{1}{4}$ from the optimal one. For the proof we extend Hartline and Roughgarden's notion of commensurate mechanisms to arbitrary single-parameter environments. Finally, for the multi-parameter environment we propose a mechanism that, using two samples, achieves on expectation over the draw of those two samples, an expected revenue within a factor of 0.589 from the expected revenue achieved in our sample complexity approach. Under the same set of distribution assumptions as in our sample complexity approach, this is a 0.589-approximation of the expected social welfare of VCG, minus some smaller additive terms.

Keywords: Automated Mechanism Design, Deferred-Acceptance Auctions, Multi-Unit Auctions, PAC Learning, t-Level Auctions, Sample Complexity.

Ευχαριστίες

Η εκπόνηση της διπλωματικής μου εργασίας σηματοδοτεί και επισήμως την ολοκλήρωση των προπτυχιακών μου σπουδών στη σχολή των Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών. Σε αυτή μου την προσπάθεια υπήρχαν πολλοί που με βοήθησαν και για αυτό θα ήθελα να τους ευχαριστήσω. Ο κ. Φωτάκης και ο κ. Παπασπύρου από την αρχή των σπουδών μου μου κίνησαν το ενδιαφέρον να ασχοληθώ περαιτέρω με το αντικείμενο της (Θεωρητικής) Πληροφορικής. Η καθοδήγηση του κ. Φωτάκη, ως επιβλέποντα καθηγητή μου, ήταν καθοριστικής σημασίας στο πρώτο μου αυτό ερευνητικό έργο. Ιδιαίτερα σημαντική για την πραγματοποίηση της παρούσας εργασίας ήταν και η συμβολή του διδακτορικού του φοιτητή Παναγιώτη Πατσιλινάκου. Η φοιτητική μου εμπειρία ήταν μοναδική χάρη στους φίλους μου που γνώρισα στη σχολή και μαζί πορευτήκαμε αυτά τα χρόνια. Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου για την απεριόριστη στήριξη και κατανόησή τους όλα αυτά τα χρόνια.

Ερμής

Contents

Chapter 1

Εκτεταμένη Ελληνική Περίληψη

Στο κεφάλαιο αυτό θα συνοψίσουμε το περιεχόμενο της παρούσας διπλωματικής, δίνοντας βασικούς ορισμούς και παρουσιάζοντας τα κυριότερα αποτελέσματά μας, χωρίς αποδείξεις

1.1 Εισαγωγή

Στη θεωρία παιγνίων, με τον όρο 'παίγνιο' εννοούμε κάθε κατάσταση όπου το τελικό αποτέλεσμα εξαρτάται από τις δράσεις δύο ή περισσότερων μελών που σκέπτονται και πράττουν στρατηγικά, οι οποίοι ονομάζονται στρατηγικοί παίκτες. Με τον όρο 'σχεδιασμός μηχανισμών' (Mechanism Design) εννοούμε το σχεδιασμό ενός παιγνίου τέτοιου ώστε, όταν οι παίκτες πράττουν στρατηγικά, το παίγνιο να οδηγείται σε μία λύση που μεγιστοποιεί κάποια αντικειμενική συνάρτηση. Στα παίγνια αυτά, κάθε παίκτης έχει τη δική του συνάρτηση 'ωφέλειας' (utility) και δρα με μοναδικό γνώμονα τη μεγιστοποίηση της δικής του ωφέλειας στο τελικό αποτέλεσμα του παιγνίου. Κοινή πρακτική στο σχεδιασμό μηχανισμών είναι να εξασφαλίζει ότι οι παίκτες δεν έχουν κίνητρο να δηλώσουν ψευδείς πληροφορίες στο μηχανισμό. Με αυτό τον τρόπο, ο σχεδιαστής έχει ισχυρότερες εγγυήσεις για το τελικό αποτέλεσμα. Οι μηχανισμοί χωρίζονται σε δύο βασικές κατηγορίες, ανάλογα με τον τρόπο που αλληλεπιδρούν οι παίκτες με αυτούς: Στους μηχανισμούς 'άμεσης αποκάλυψης' (direct revelation), στους οποίους οι παίκτες ανακοινώνουν άμεσα στο μηχανισμό τις προτιμήσεις τους και στους 'έμμεσους', όπου οι παίκτες δε δηλώνουν τις προτιμήσεις τους στο μηχανισμό, αλλά ο μηχανισμός εξελίσσεται ανάλογα με τις πράξεις των παικτών. Ένας direct revelation μηχανισμός με n παίκτες και ένα σύνολο $\mathcal{O}$ από δυνατά αποτελέσματα λειτουργεί ως εξής: Κάθε παίκτης i έχει μία 'συνάρτηση αξίας' (valuation function) $v_i : \mathcal{O} \to \mathcal{R}_+$ και κάνει στο μηχανισμό μία 'δήλωση' (bid) της αξίας του για κάθε αποτέλεσμα, b_i. Ο μηχανισμός αρχικά συλλέγει τα bids όλων των παικτών, και έπειτα αντιστοιχίζει το $\boldsymbol{b} = (b_1, b_2, \ldots, b_n) \in B$ σε ένα πιθανό αποτέλεσμα $O \in \mathcal{O}$ μέσω μίας συνάρτησης $f : B \to \mathcal{O}$. Η συνάρτηση αυτή συχνά αναφέρεται στη βιβλιογραφία ως 'allocation function'. Τέλος, ο μηχανισμός χρησιμοποιεί μία δεύτερη συνάρτηση $p : B \to R_+^n$ για να υπολογίσει το διάνυσμα πληρωμών των παικτών.

Σε ένα single-parameter περιβάλλον κάθε παίκτης ανακοινώνει στο μηχανισμό μία μόνο τιμή, ενώ σε multi-parameter πολλές.

Συνήθως ένα **single-parameter** περιβάλλον ορίζεται ως εξής:

- n στρατηγικοί παίκτες.

- Ένα πεπερασμένο σύνολο $\mathcal{O}$ από εφικτά αποτελέσματα. Σε κάθε αποτέλεσμα ένα υποσύνολο των παικτών 'κερδίζουν', ενώ οι υπόλοιποι όχι.
- Κάθε παίκτης έχει μία αξία, η οποία είναι κοινή, για όλα τα αποτελέσματα στα οποία κερδίζει, δηλαδή δεν εξαρτάται από το ποιοι άλλοι παίκτες 'κέρδισαν'. Αυτή του ακριβώς την αξία δηλώνει στο μηχανισμό.

Αντίστοιχα, ένα **multi-parameter** περιβάλλον ορίζεται ως εξής:

- n στρατηγικοί παίκτες.
- Ένα πεπερασμένο σύνολο $\mathcal{O}$ από εφικτά αποτελέσματα.
- Κάθε παίκτης i έχει μία προσωπική αξία $v_i(O)\ \forall O \in \mathcal{O}$.

Η *ωφέλεια* (utility) κάθε παίκτη i ορίζεται συνήθως ως η διαφορά της αξίας του για το αποτέλεσμα που καθόρισε ο μηχανισμός, μείον το τι ο παίκτης αυτός καλείται να πληρώσει από το μηχανισμό, δηλαδή $u_i(\boldsymbol{b}) = v_i(f(\boldsymbol{b})) - p_i(\boldsymbol{b})$. Ωφέλειες αυτής της μορφής ονομάζονται στη βιβλιογραφία 'ψευδο-γραμμικές' (quasi-linear). Η στρατηγική b_i του παίκτη i ονομάζεται *κυρίαρχη στρατηγική* εάν μεγιστοποιεί την ωφέλειά του, ανεξαρτήτως των στρατηγικών $\boldsymbol{b}_{-i}$ των υπολοίπων παικτών. Τέλος, ένας μηχανισμός ονομάζεται *φιλαλήθης* εάν για κάθε παίκτη είναι κυρίαρχη στρατηγική να ανακοινώσει στο μηχανισμό την πραγματική του αξία για κάθε αποτέλεσμα.
Για single-parameter περιβάλλοντα, ο Myerson προσδιόρισε την ικανή και αναγκαία συνθήκη ώστε να είναι ένας μηχανισμός φιλαλήθης και ταυτόχρονα κανένας παίκτης να μην έχει αρνητική ωφέλεια, εάν πάρει μέρος στο μηχανισμό. Οι 2 αυτές συνθήκες μαζί στη βιβλιογραφία ονομάζονται *Dominant Strategy Incentive Compatitable* (DSIC). Σύμφωνα με το *Λήμμα του Myerson*, ένας μηχανισμός με allocation function f και payment rule p για ένα single-parameter περιβάλλον είναι DSIC εάν και μόνο αν για κάθε παίκτη i με δήλωση b_i στο μηχανισμό και δηλώσεις $\boldsymbol{b}_{-i}$ από τους υπόλοιπους παίκτες, ισχύει:

- Η $f_i(b_i, \boldsymbol{b}_{-i})$ είναι αύξουσα στο πρώτο της όρισμα.
- Υπάρχει ένας μοναδικός κανόνας πληρωμών που καθιστά το μηχανισμό (f, p) DSIC. Αυτός δίνεται από τον τύπο:

$$p_i(b_i, \boldsymbol{b}_{-i}) = \int_0^{b_i} z \frac{d}{dz} f_i(z, \boldsymbol{b}_{-i}) dz \tag{1.1}$$

Οι *δημοπρασίες k-μονάδων* (k-unit auctions) αποτελούν μία ειδική υποκατηγορία των multi-parameter περιβαλλόντων. Σε αυτές, υπάρχουν k, όπου $k \in \mathbb{N} \cup \{\infty\}$ διαθέσιμα 'αντίτυπα' του ίδιου αγαθού διαθέσιμα προς πώληση. Κάθε παίκτης έχει ένα valuation function $v_i : \mathbb{N} \to \mathbb{R}_+$, δηλαδή η αξία του εξαρτάται από το πόσα αντίτυπα του αγαθού θα πάρει. Πιθανός σκοπός του δημοπράτη στο περιβάλλον αυτό, όπως το ορίσαμε, θα μπορούσε να είναι να υπολογίσει μία ανάθεση $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ τέτοια ώστε να μεγιστοποιείται το άθροισμα των αξιών όλων των παικτών για το τελικό αποτέλεσμα. Το άθροισμα αυτό συναντάται στη βιβλιογραφία ως 'κοινωνική ωφέλεια' (Social Welfare) και ορίζεται ως $\sum_{i=1}^{n} v_i(x_i)$. Ένας άλλος πιθανός σκοπός του δημοπράτη στο παίγνιο αυτό θα μπορούσε να ήταν να μεγιστοποιήσει το *εισόδημά του* από τη δημοπρασία, δηλαδή το άθροισμα των πληρωμών όλων των παικτών στο τελικό αποτέλεσμα του μηχανισμού. Η κοινωνική ωφέλεια και το εισόδημα στο τελικό αποτέλεσμα είναι οι δύο πιο συχνές αντικειμενικές συναρτήσεις στο σχεδιασμό μηχανισμών.
Ένας από τους μηχανισμούς που συναντούνται πιο συχνά στη βιβλιογραφία είναι ο VCG, ο οποίος παρουσιάζεται αναλυτικά στην υποενότητα 3.3.3. Ο λόγος για τον οποίο συναντάται

τόσο συχνά στη βιβλιογραφία είναι διότι είναι φιλαλήθης και ταυτόχρονα μεγιστοποιεί, σε κάθε multi-parameter περιβάλλον και για κάθε *valuation profile* $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ την κοινωνική ωφέλεια. Ωστόσο, ο VCG μηχανισμός δεν εφαρμόζεται συχνά στην πράξη. Ο λόγος είναι ότι σε πολλά περιβάλλοντα, η υπολογιστική του πολυπλοκότητα είναι τόσο μεγάλη που καθιστά τη χρήση του αδύνατη.

Η κοινωνική ωφέλεια είναι, κατά μία έννοια, μοναδική: Για κάθε άλλη αντικειμενική συνάρτηση που συναντάται στην πράξη, δεν υπάρχει φιλαλήθης μηχανισμός που να τη μεγιστοποιεί σε κάθε περιβάλλον, για κάθε valuation profile. Συνήθως λοιπόν, σκοπός του σχεδιαστή μηχανισμών είναι να προσδιορίσει έναν μηχανισμό που μεγιστοποιήσει την *αναμενόμενη τιμή* της εκάστοτε αντικειμενικής συνάρτησης.

Το ίδιο ισχύει και για το εισόδημα. Ας πάρουμε την περίπτωση όπου υπάρχει μόνο ένα αγαθό προς πώληση και ένας υποψήφιος αγοραστής με προσωπική αξία v για το αγαθό. Με μόνο έναν υποψήφιο αγοραστή, ο χώρος των direct-revelation DSIC μηχανισμών για το πρόβλημα αυτό είναι πολύ μικρός: είναι ακριβώς όλες οι πιθανές 'αναρτημένες τιμές' (posted prices), δηλαδή ο μηχανισμός προσφέρει το αγαθό σε κάποια τιμή και ο υποψήφιος αγοραστής μπορεί είτε να την δεχτεί, και να αγοράσει το αγαθό στην τιμή αυτή, είτε να την απορρίψει. Το να μεγιστοποιήσει κανείς την κοινωνική ωφέλεια σε αυτό το περιβάλλον είναι τετριμμένο: Θέτοντας τιμή μηδέν για το αγαθό, ο υποψήφιος αγοραστής θα παίρνει πάντα το αγαθό, αφού είναι δωρεάν. Ας υποθέσουμε όμως ότι ο σχεδιαστής μηχανισμών ήθελε να μεγιστοποιήσει το εισόδημα του μηχανισμού. Πώς έπρεπε να θέσει την τιμή για το αγαθό; Εκ των υστέρων (με άλλα λόγια, εάν μπορούσε να μαντέψει την αξία v του υποψήφιου αγοραστή για το αγαθό), θα έπρεπε να θέσει την τιμή r του αγαθού ίση με v. Μη γνωρίζοντας την τιμή v, είναι προφανές ότι δεν υπάρχει posted-price μηχανισμός που να μεγιστοποιεί το εισόδημα για κάθε πιθανή αξία v του υποψηφίου αγοραστή. Για το λόγο αυτό, όταν η αντικειμενική συνάρτηση που ενδιαφέρει τον δημοπράτη είναι το εισόδημα του μηχανισμού, σκοπός του σχεδιαστή μηχανισμών είναι να μεγιστοποιήσει το *αναμενόμενο εισόδημα* του μηχανισμού.

Υπάρχει όμως DSIC μηχανισμός που να μεγιστοποιεί το αναμενόμενο εισόδημα, δεδομένων των κατανομών των αξιών των παικτών για τα διαφορετικά ενδεχόμενα; Η απάντηση, για single-parameter περιβάλλοντα, είναι θετική. Ο μηχανισμός αυτός προκύπτει από την παρατήρηση ότι για single-parameter περιβάλλοντα, οι πληρωμές των παικτών δίνονται από το λήμμα του Myerson. Παίρνοντας αναμενόμενες τιμές πάνω στις πληρωμές αυτές και μεγιστοποιώντας την ανάθεση του μηχανισμού ως προς αυτές προκύπτει ο αντίστοιχος μηχανισμός, ο οποίος ονομάζεται επίσης Myerson μηχανισμός. Η μόνη προϋπόθεση είναι ότι οι κατανομές των παικτών είναι *κανονικές*, το οποίο διαισθητικά σημαίνει ότι οι 'ουρές τους' δεν πρέπει να φθίνουν πάρα πολύ αργά.

Η επόμενη λογική ερώτηση είναι εάν υπάρχει κάτι αντίστοιχο με το μηχανισμό Myerson για multi-parameter περιβάλλοντα, δηλαδή ένας DSIC μηχανισμός που, δεδομένων των κατανομών των αξιών των παικτών, να μεγιστοποιεί το αναμενόμενο εισόδημα. Η απάντηση είναι αρνητική: Ακόμα και για την φαινομενικά απλή περίπτωση όπου υπάρχουν μόνο 2 αγαθά διαθέσιμα προς πώληση και 2 υποψήφιοι αγοραστές, ο μηχανισμός που μεγιστοποιεί το αναμενόμενο εισόδημα δεν είναι γνωστός.

1.2 Δημοπρασίες Αναβαλλόμενης Αποδοχής

Οι δημοπρασίες αναβαλλόμενης αποδοχής (deferred-acceptance auctions / DA auctions) προτάθηκαν από τους Milgrom και Segal. Στην αρχική τους μορφή, αυτή η οικογένεια ήταν περιορισμένη σε single-parameter περιβάλλοντα, όπου κάθε παίκτης στο τελικό αποτέλεσμα είτε θα είναι 'νικητής', που σημαίνει ότι θα εξυπηρετηθεί από το μηχανισμό, είτε θα είναι 'χαμένος'. Όπως είναι συνηθισμένο σε τέτοια περιβάλλοντα, στο μηχανισμό κάθε παίκτης ανακοινώνει την αξία που έχει για αυτόν να νικήσει, ενώ ο μηχανισμός θεωρεί ότι κάθε παίκτης έχει μηδενική αξία για τα αποτελέσματα στα οποία χάνει.

Μία δημοπρασία αναβαλλόμενης αποδοχής εξελίσσεται με γύρους. Αρχικά, όλοι οι παίκτες είναι ενεργοί. Σε κάθε γύρο, ο παίκτης που φαίνεται 'λιγότερο υποσχόμενος' αποκλείεται από το μηχανισμό, μέχρις ότου ου παίκτες που απομένουν να αποτελούν εφικτή λύση. Οι πληρωμές προκύπτουν απευθείας από το λήμμα του Myerson: Κάθε νικητής πληρώνει το λιγότερο που θα μπορούσε να δηλώσει στο μηχανισμό, δεδομένων των δηλώσεων των υπολοίπων παικτών και να εξυπηρετηθεί από το μηχανισμό. Από αυτοί τους την ιδιαιτερότητα προκύπτει και το όνομα των δημοπρασιών αναβαλλόμενης αποδοχής: Σε αντίθεση με το συνηθισμένο, ο μηχανισμός αποφασίζει με (εν δυνάμει δυναμικό) άπληστο τρόπο για το ποιους παίκτες θα αποκλείσει, όχι ποιους παίκτες θα αποδεχτεί, μέχρις ότου το σύνολο των εναπομείναντων παικτών αν αποτελούν εφικτή λύση.

Μία λεπτομέρεια ζωτικής σημασίας για τις δημοπρασίες αναβαλλόμενης αποδοχής είναι ότι ο μηχανισμός, όταν αξιολογεί τον εκάστοτε παίκτη, δε λαμβάνει υπόψιν του τις δηλώσεις των υπόλοιπων παικτών που δεν έχουν αποκλειστεί ακόμη. Αυτό έχει ως αποτέλεσμα ότι κάθε νικητής δε μπορεί αλλάζοντας την προσφορά του να επηρεάσει της πληρωμές των υπολοίπων νικητών, εκτός και εάν αλλάξει την προσφορά του σε κάποια προσφορά με την οποία θα έχανε ο ίδιος.

Ο λόγος για τον οποίο οι Milgrom και Segal πρότειναν την οικογένεια αυτή μηχανισμών είναι γιατί όπως απέδειξαν, οι δημοπρασίες αναβαλλόμενης αποδοχής παρουσιάζουν ορισμένες αξιοσημείωτες, και σπάνιες, ιδιότητες. Αρχικά, οι δημοπρασίες αναβαλλόμενης αποδοχής είναι *προφανώς φιλαλή-/θεις* (obviously strategyproof - OSP). Μία στρατηγική λέγεται *προφανώς κυρίαρχη* (obviously dominant) όταν, για κάθε πιθανή διαφοροποίηση από αυτήν, στο σημείο όπου οι δύο στρατηγικές διαφοροποιούνται για πρώτη φορά και με τις πληροφορίες που διαθέτει ο παίκτης στο σημείο αυτό, ακόμα και το καλύτερο δυνατό αποτέλεσμα υπό την διαφοροποίηση δεν είναι καλύτερο από το χειρότερο δυνατό αποτέλεσμα για τον παίκτη υπό την κυρίαρχη στρατηγική. Αυτή η ιδιότητα μπορεί να ερμηνευθεί και από την οπτική γωνία των παικτών: Μία στρατηγική είναι προφανώς κυρίαρχη αν και μόνο αν ακόμα και ένας "νοητικά περιορισμένος" παίκτης μπορεί να συνειδητοποιήσει ότι είναι ασθενώς κυρίαρχη. Ένας μηχανισμός ονομάζεται *προφανώς φιλαλήθης* εάν υπάρχει σημείο ισορροπίας με μόνο προφανώς κυρίαρχες στρατηγικές. Στην πράξη, το να είναι ένας μηχανισμός OSP τον καθιστά 'προσεγγίσιμο' και από παίκτες που δεν κατανοούν αλγοριθμική θεωρία παιγνίων. Αυτό είναι σημαντικό σε περιπτώσεις όπου η συμμετοχή περισσότερων παικτών στο παίγνιο βελτιώνει την ποιότητα του τελικού αποτελέσματος.

Μία επιπλέον αξιοσημείωτη ιδιότητα των δημοπρασιών αναβαλλόμενης αποδοχής είναι ότι είναι ασθενώς *γκρουπ-φιλαλήθεις* (weakly group strategy-proof - WGSP). Αυτό σημαίνει ότι κανένας συνασπισμός παικτών δε μπορεί να συνωμοτήσει με τέτοιο τρόπο κατά του δημοπράτη και συλλογικά να υποβάλλουν ψευδείς δηλώσεις τέτοιες ώστε όλα τα μέλη του συνασπισμού να έχουν μεγαλύτερη ωφέλεια από ότι εάν δεν είχαν συμμετάσχει στο συνασπισμό και αντί αυτού είχαν δηλώσει την πραγματική τους αξία. Στην πράξη, αυτό συνεπάγεται ότι κανένας παίκτης δεν έχει κίνητρο να συνασπιστεί με άλλους κατά του δημοπράτη. Στο σύγχρονο κόσμο, η υπόθεση ότι οι παίκτες σε ένα παιχνίδι δε θα επικοινωνήσουν μεταξύ τους, ιδιαίτερα σε αγορές υψηλής σημασίας με λίγους παίκτες, είναι μη ρεαλιστική.

Οι Γκατζέλης, Μαρκάκης και Roughgarden επέκτειναν την έννοια των δημοπρασιών αναβαλλόμενης αποδοχής σε περιβάλλοντα όπου υπάρχουν πολλαπλά "επίπεδα εξυπηρέτησης" από το μηχανισμό, εξασφαλίζοντας ότι όλες οι σημαντικές ιδιότητες των δημοπρασιών αναβαλλόμενης αποδοχής, όπως είχαν οριστεί από τους Milgrom και Segal, διατηρούνται. Αυτή τη νέα κατηγορία μηχανισμών την ονόμασαν *γενικευμένες δημοπρασίες αναβαλλόμενης αποδοχής* (generalized single-parameter DA auction). Συγκεκριμένα:

Ορισμός (Γενικευμένες δημοπρασίες αναβαλλόμενης αποδοχής). Μία γενικευμένη δημοπρασία αναβαλλόμενης αποδοχής λειτουργεί σε διακεκριμένους γύρους $t \geq 1$. Συμβολίζουμε με $A_t \subseteq \mathcal{N}$

το σύνολο των ακόμα-ενεργών παικτών στην αρχή του γύρου t. Αρχικά είναι $A_1 = \mathcal{N}$, και $A_{t+1} \subset A_t$ για κάθε $t \geq 1$. Η δημοπρασία αναβαλλόμενης αποδοχής καθορίζεται πλήρως από 2 σύνολα συναρτήσεων:

- Τις **συναρτήσεις αξιολόγησης** $\sigma_i^{A_t}(b_i, \boldsymbol{b}_{\mathcal{N}\setminus A_t})$, οι οποίες είναι ασθενώς αύξουσες στο πρώτο τους όρισμα.
- Τις **συναρτήσεις 'εξασφάλισης'** (clinching functions) $g_i^{A_t}(\boldsymbol{b}_{\mathcal{N}\setminus A_t})$, οι οποίες είναι μη αύξουσες ως προς το σύνολο των ενεργών παικτών, δηλαδή $g_i^{A_{t+1}}(\boldsymbol{b}_{\mathcal{N}\setminus A_{t+1}}) \geq g_i^{A_t}(\boldsymbol{b}_{\mathcal{N}\setminus A_t})$.

Σε κάθε γύρο t, εάν $A_t \neq \emptyset$, τότε κάποιος ενεργού παίκτη $i \in \arg\min_{i \in A_t}\{\sigma_i^{A_t}(b_i, \boldsymbol{b}_{N\setminus A_t})\}$ αποκλείεται από το μηχανισμό, το επίπεδο εξυπηρέτησής του οριστικοποιείται σύμφωνα με τη συνάρτηση εξασφάλισής του σε $g_i^{A_t}(\boldsymbol{b}_{\mathcal{N}\setminus A_t})$ και το επόμενο σύνολο ενεργών παικτών είναι το $A_{t+1} = A_t \setminus \{i\}$.
Η επίδοση των μηχανισμών αναβαλλόμενες αποδοχής έχουν μελετηθεί εκτενώς, κυρίως όμως από την οπτική γωνία της worst-case analysis. Σε αυτή τη διπλωματική εργασία προσπαθούμε, κάνοντας κάποιες εύλογες υποθέσεις για τις κατανομές των αξιών των παικτών, να εντοπίσουμε περιβάλλοντα όπου μπορεί κανείς, χρησιμοποιώντας κάποια δείγματα από τις αξίες των παικτών, να κατασκευάσει δημοπρασίες αναβαλλόμενης αποδοχής με υψηλό αναμενόμενο εισόδημα. Έπειτα, μελετάμε πώς τροποποιούνται αυτά τα αποτελέσματα εάν περιορίσουμε σημαντικά τον αριθμό των δειγμάτων.

1.3 Εκμάθηση Δημοπρασιών Αναβαλλόμενης Αποδοχής

Σε αυτή την ενότητα θα μελετήσουμε το πρόβλημα της εκμάθησης δημοπρασιών με υψηλό αναμενόμενο εισόδημα για δημοπρασίες k-αντιτύπων του ίδιου αγαθού. Η ενότητα χωρίζεται σε 2 υποενότητες, με βάση το εάν βρισκόμαστε σε single- ή multi-parameter περιβάλλον.

1.3.1 Single-Parameter Περιβάλλοντα

Στην υποενότητα αυτή θα προτείνουμε μία νέα κλάση δημοπρασιών αναβαλλόμενης αποδοχής και θα θέσουμε ένα άνω φράγμα στην δειγματική πολυπλοκότητα που απαιτείται ώστε να προσδιοριστεί μία δημοπρασία μέσα από την οικογένεια αυτή με αναμενόμενο εισόδημα αυθαίρετα κοντά στο βέλτιστο για single-parameter multi-unit auctions. Κάθε παίκτης κατά τη διάρκεια της δημοπρασίας ανακοινώνει ένα νούμερο στο μηχανισμό: την αξία του ανά αντίτυπο του αγαθού, ενώ ο μηχανισμός γνωρίζει από πριν μέχρι τον μέγιστο αριθμό αντικειμένων για τα οποία ενδιαφέρεται ο κάθε παίκτης. Το περιβάλλον ορίζεται τυπικά ως εξής:

Το Περιβάλλον (Παίκτες με αθροιστικές αξίες)

- n παίκτες.
- Υπάρχουν m αντίτυπα του αγαθού διαθέσιμα προς πώληση.
- Κάθε παίκτης i έχει μία γνωστή εκ των προτέρων στο μηχανισμό *ζήτηση* d_i.
- Κάθε παίκτης i έχει αθροιστική αξία, μέχρι τη ζήτησή του d_i:
  Για κάθε παίκτη i, η αξία του για x_i αντίτυπα του αγαθού, δεδομένου ότι $x_i \leq d_i$, είναι $x_i \cdot v_i$.
- Η αξία του κάθε παίκτη i ανά αντίτυπο του αγαθού, v_i, ακολουθεί κάποια κατανομή με συνάρτηση πυκνότητας πιθανότητας $f(\cdot)$ και virtual valuation function $\phi_i(\cdot)$.

- Οι αξίες και των n παικτών ανά αντίτυπο του αγαθού είναι φραγμένες στο $[1, H]$.

Η οικογένεια μηχανισμών που προτείνουμε για το περιβάλλον αυτό ήταν έντονα εμπνευσμένη από την οικογένεια των t-level auctions των Morgenstern και Roughgarden ([23]). Για το λόγο αυτό την ονομάσαμε “linear DA t-level auctions”.

Ορισμός 5.4 (Γραμμικές DA δημοπρασίες t-επιπέδων). Κάθε παίκτης i αντιμετωπίζει t κατώφλια: $0 \leq l_{i,0} \leq l_{i,1} \leq \cdots \leq l_{i,t-1}$. Αυτό το σύνολο $t \cdot n$ αριθμών ορίζει μία γραμμική δημοπρασία αναβαλλόμενης αποδοχής t-επιπέδων με τον παρακάτω κανόνα ανάθεσης: Έστω $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ το διάνυσμα των αξιών (valuation profile) των παικτών:

1. Για κάθε παίκτη i με $t_i(v_i)$ συμβολίζουμε τον δείκτη τ του μεγαλύτερου του κατωφλίου $l_{i,\tau}$ που είναι μικρότερο ή ίσο του v_i (ή -1, εάν $v_i < l_{i,0}$). Το $t_i(v_i)$ το αποκαλούμε το *επίπεδο* του παίκτη i.
2. Ταξινόμησε όλους τους παίκτες από το χαμηλότερο επίπεδο προς το υψηλότερο, και για τους παίκτες που βρίσκονται στο ίδιο επίπεδο χρησιμοποίησε έναν ντετερμινιστικό κανόνα $\prec$ για να διαλέγεις νικητή (ποιος είναι ψηλότερα από ποιον εντός του ίδιου επιπέδου).
3. **DA ανάθεση**: Ξεκίνα να αποκλείεις τους παίκτες, από το χαμηλότερο επίπεδο προς το υψηλότερο. Κάθε παίκτης, τη στιγμή που τερματίζεται κερδίζει, μέχρι το όριο της ζήτησής του, όσο το δυνατόν περισσότερα αντικείμενα, δεδομένου ότι τα εναπομείναντα αντικείμενα αρκούν για να καλύψουν πλήρως τη ζήτηση των ακόμα-ενεργών παικτών.
4. **Κανόνας πληρωμών**: Ο μοναδικός που καθιστά το μηχανισμό φιλαλήθη. Κάθε παίκτης, για κάθε αντίτυπο που κερδίζει, πληρώνει το ελάχιστο που θα μπορούσε να είχε δηλώσει στο μηχανισμό και να κερδίζει το αντικείμενο αυτό.

Πρόταση: Η οικογένεια δημοπρασιών που μόλις περιγράφηκε περιέχει μόνο έγκυρες δημοπρασίες αναβαλλόμενης αποδοχής. Δεδομένων των κατωφλιών τα αντίστοιχα scoring και clinching functions είναι:

- scoring functions: $\sigma_i^{A_t}(b_i, \boldsymbol{b}_{N \setminus A_t}) = t_i(v_i)$.
- clinching functions: $g_i^{A_t}(b_{N \setminus A_t}) = \min \left\{ d_i, m - \min \left\{ m, \sum_{j \in A_t \setminus \{i\}} d_j \right\} \right\}$.

Για αυτή την υποκατηγορία των δημοπρασιών αναβαλλόμενης αποδοχής που μόλις ορίσαμε υπολογίσαμε την δειγματική πολυπλοκότητα ώστε να προσδιορίσει κανείς, για κάθε single-parameter περιβάλλον όπως περιγράφηκε στην αρχή της υποενότητας, μία δημοπρασία από την οικογένεια με αναμενόμενο εισόδημα αυθαίρετα κοντά στο βέλτιστο, δηλαδή το αναμενόμενο εισόδημα του μηχανισμού Myerson. Η διαδικασία που ακολουθήσαμε είναι αρκετά συνηθισμένη τα τελευταία χρόνια στην αυτοματοποιημένη σχεδίαση μηχανισμών. Αρχικά φράξαμε το *generalizaiton error* του μηχανισμού, δηλαδή για κάθε μηχανισμό από την οικογένεια, τη μέγιστη απόκλιση που μπορεί να έχει το αναμενόμενο εισόδημά του στον πραγματικό κόσμο σε σχέση με το *εμπειρικό του εισόδημα* στα δείγματα. Για να το επιτύχουμε αυτό, χρειάστηκε επιπλέον να φράξουμε την ψευδο-διάσταση της οικογένειας. Έπειτα φράξαμε το *representation error* του μηχανισμού, δηλαδή το πόσο πολύ μπορεί να προσεγγίσει το εισόδημα του βέλτιστου μηχανισμού από την οικογένεια το βέλτιστο εισόδημα του μηχανισμού Myerson. Συνδυάζοντας τα generalization και representation error bounds που δείξαμε και λύνοντας ως προς τον αριθμό δειγμάτων, προκύπτει ένα άνω φράγμα για τη δειγματική πολυπλοκότητα της οικογένειας.

Θεώρημα 5.2. Η ψευδο-διάσταση των γραμμικών δημοπρασιών αναβαλλόμενης αποδοχής t-επιπέδων είναι $O(nt\log(nt))$.

Θεώρημα 5.3. Στο περιβάλλον που περιγράφηκε στην αρχή της υποενότητας αυτής, για $t = \lceil \frac{2}{\epsilon} \rceil + \log_{1+\frac{\epsilon}{2}} H = \Theta\left(\frac{1}{\epsilon} + \log_{1+\epsilon} H\right)$ η κλάση των γραμμικών δημοπρασιών αναβαλλόμενης αποδοχής t-επιπέδων περιέχει μία δημοπρασία με αναμενόμενο εισόδημα τουλάχιστον $(1-\epsilon)$ φορές το βέλτιστο.

Οι αποδείξεις και των 2 αυτών θεωρημάτων είναι αρκετά τεχνικές, και για αυτό το λόγο παραλείπονται από την περίληψη αυτή. Περιέχονται όμως στο κύριο σώμα της διπλωματικής. Συνδυάζοντάς τα 2 αυτά θεωρήματα, αρκετά εύκολα μπορεί να αποδείξει κανείς το παρακάτω πόρισμα:

Πόρισμα 5.2. Με πιθανότητα τουλάχιστον $1-\delta$, ο μηχανισμός από την κλάση των γραμμικών δημοπρασιών αναβαλλόμενης αποδοχής t-επιπέδων που εμπειρικά μεγιστοποιεί το εισόδημα σε ένα σύνολο N δειγμάτων είναι μία $1-O(\epsilon)$-προσσέγιση του Myerson για m αντίτυπα προς πώληση και n παίκτες με αξίες στο $[1, H]$, για $t = O(\frac{1}{\epsilon} + \log_{1+\epsilon} H)$ και

$$N = O\left(\left(\frac{Hm}{\epsilon}\right)^2 \left(nt\log(nt)\ln\frac{Hm}{\epsilon} + \ln\frac{1}{\delta}\right)\right) = \widetilde{O}\left(\frac{H^2m^2n}{\epsilon^3}\right). \tag{1.2}$$

Το αποτέλεσμα αυτό ήταν, κατά μία έννοια, αναμενόμενο. Οι Γκατζέλης, Μαρκάκης και Roughgarden έδειξαν ότι σε single-parameter περιβάλλοντα αυτής της μορφής ο VCG μηχανισμός, ο οποίος ορίζεται ως ο μηχανισμός που μεγιστοποιεί την κοινωνική ωφέλεια, μπορεί να υλοποιηθεί ως γενικευμένη δημοπρασία αναβαλλόμενης αποδοχής. Ο Myerson μηχανισμός μεγιστοποιεί το αναμενόμενο εισόδημα, ή ισοδύναμα μεγιστοποιεί την εικονική κοινωνική ωφέλεια, η οποία είναι ίση με τις αναμενόμενες πληρωμές των παικτών. Εφόσον λοιπόν ο μηχανισμός που μεγιστοποιεί την κοινωνική ωφέλεια είναι υλοποιήσιμος ως δημοπρασία αναβαλλόμενης αποδοχής, το ίδιο πρέπει να ισχύει και για τον μηχανισμό που μεγιστοποιεί την εικονική κοινωνική ωφέλεια, δηλαδή για τον Myerson. Επομένως το πρόβλημα καταλήγει στο να προσδιορίσει κανείς την υλοποίηση αυτή του Myerson και έπειτα να φράξει τον απαιτόμενο αριθμό δειγμάτων ώστε να μάθει ο μηχανισμός τις συναρτήσεις εικονικής αξίας (virtual valuation functions) όλων των παικτών.

1.3.2 Multi-Parameter Περιβάλλοντα

Στην υποενότητα αυτή θα προτείνουμε μία κλάση δημοπρασιών, οι οποίες είναι τετριμμένα αναβαλλόμενης αποδοχής, και θα θέσουμε ένα κάτω φράγμα στο αναμενόμενο εισόδημα της εμπειρικά βέλτιστης δημοπρασίας από την οικογένεια αυτή, συναρτήσει του αριθμού των δειγμάτων που χρησιμοποιήθηκαν για τον προσδιορισμό της. Τώρα βρισκόμαστε σε multi-parameter περιβάλλον: κάθε παίκτης κατά τη διάρκεια της δημοπρασίας ανακοινώνει m νούμερα στο μηχανισμό: την αξία του για το ένα αντίτυπο του αγαθού, έπειτα την επιπλέον αξία του για ένα δεύτερο, και ούτω καθεξής. Το περιβάλλον ορίζεται τυπικά ως εξής:

Το Περιβάλλον (Παίκτες με submodular αξίες)

- n παίκτες.
- Υπάρχουν m αντίτυπα του αγαθού διαθέσιμα προς πώληση.
- Κάθε παίκτης i δηλώνει τις m *οριακές του αξίες* στο μηχανισμό, $v_{i,1}, v_{i,2}, \ldots, v_{i,m}$, όπου $v_{i,j}$: Η αξία που έχει για τον παίκτη i να αποκτήσει το j-οστό αντίτυπό του του

αγαθού, δεδομένου ότι έχει εξασφαλίσει ήδη $j-1$ αντίτυπα. Για κάθε παίκτη i, η αξία του για x_i αντίτυπα του αγαθού, δεδομένου ότι $x_i \leq d_i$, είναι $x_i \cdot v_i$.

- Η συνάρτηση αξίας κάθε παίκτη είναι submodular: $v_{i,1} \geq v_{i,2} \geq \cdots \geq v_{i,m} \; \forall i$
- Κάθε οριακή αξία $v_{i,j}$ ακολουθεί κάποια Gaussian ή Sub-Gaussian κατανομή $F_{i,j}$.

Αυτό το σύνολο υποθέσεων είναι αρκετά φυσικό: Ένα μεγάλο εύρος κατανομών ανήκουν στις Sub-Gaussian, συμπεριλαμβανομένων των Gaussian και των mixtures of Gaussians, των ομοιόμορφων και κάθε άλλης φραγμένης κατανομής. Διαισθητικά, κάθε κατανομή με ουρές οι οποίες φθίνουν τουλάχιστον τόσο γρήγορα όσο αυτές κάποιας Gaussian κατανομήμής είναι Sub-Gaussian.
Αυτό το περιβάλλον είναι αρκετά πιο απαιτητικό από εκείνο της προηγούμενης υποενότητας. Τώρα δεν πρόκειται για ένα single-parameter περιβάλλον, οπότε δεν υπάρχει ο μηχανισμός Myerson ώστε να καθορίζει το βέλτιστο αναμενόμενο εισόδημα. Για το λόγο αυτό, θα πρέπει να συγκρίνουμε το αναμενόμενο εισόδημα του μηχανισμού που προτείνουμε απευθείας με την αναμενόμενη κοινωνική ωφέλεια του VCG, η οποία αποτελεί τετριμμένο άνω φράγμα για το αναμενόμενο εισόδημα οποιουδήποτε φιλαλήθη μηχανισμού.

Ορισμός 6.1 (Unit-Bundling Δημοπριασίες). Κάθε unit-bundling δημοπρασία αναβαλλόμενης αποδοχής για n παίκτες μπορεί να προσδιοριστεί μοναδικά από τις παραμέτρους της $\{(s_1, s_2, \ldots, s_n) \in \mathbb{N}_+^n \mid s_1 + s_2 + \cdots + s_n \leq m\}$ και $r_1, r_2, \ldots, r_n \in \mathbb{R}_+$. Αυτό το σύνολο $2n$ αριθμών ορίζει μία unit-bundling δημοπρασία αναβαλλόμενης αποδοχής με τους παρακάτω κανόνες ανάθεσης και πληρωμής:

- Πρόσφερε σε κάθε παίκτη i ένα bundle s_i αντιτύπων σε τιμή r_i ανά αντίτυπο, δηλαδή $r_i \cdots s_i$ συνολικά.
- Οι παίκτες που δέχονται την προσφορά τους ανατίθενται τον αντίστοιχο αριθμό αντικειμένων που όριζε το bundle τους και πληρώνουν την αντίστοιχη τιμή, ενώ οι υπόλοιποι δεν διατίθενται κανένα αντίτυπο του αντικειμένου και δεν πληρώνουν τίποτα.

Σε πρώτη ανάγνωση, η κλάση αυτή μηχανισμών ίσως να δείχνει απλοϊκή. Αυτό όμως δεν ισχύει στην πραγματικότητα. Διαισθητικά, οι δημοπρασίες αναβαλλόμενης αποδοχής 'δουλεύουν καλά' όταν οι συναρτήσεις αξίας των παικτών είναι αθροιστικές, αλλά όχι τόσο καλά όταν αυτές είναι submodular. Με την οικογένεια αυτή ο σκοπός μας είναι να προσδιορίσουμε για κάθε παίκτη, με βάση τα δείγματα, το βέλτιστο σημείο στο οποίο να 'γραμμικοποιήσουμε' τη συνάρτηση αξίας του.
Η προσέγγισή μας στη συνέχεια είναι παρεμφερής με εκείνη που ακολουθήσαμε για τις γραμμικές δημοπρασίες αναβαλλόμενης αποδοχής t-επιπέδων: Θα εδραιώσουμε ένα representation και ένα geranlization error φράγμα για την οικογένεια, και έπειτα θα συνδυάσουμε τα 2. Για το generalization error φράγμα θα χρησιμοποιήσουμε την ίδια λογική με την προηγούμενη υποενότητα, δηλαδή θα δείξουμε πρώτα ένα άνω φράγμα για την ψευδο-διάσταση της κλάσης. Για το representation error φράγμα, θα πρέπει τώρα να δείξουμε ότι η κλάση που προτείνουμε περιέχει ένα μηχανισμό με αναμενόμενο εισόδημα 'κοντά' στην αναμενόμενη κοινωνική ωφέλεια του VCG.
Για να φράξουμε την ψευδο-διάσταση της κλάσης των Unit-Bundling δημοπρασιών θα χρησιμοποιήσουμε το κύριο αποτέλεσμα από την εργασία "A General Theory of Sample Complexity for Multi-Item Profit Maximization". Για λόγους πληρότητας, το παραθέτουμε παρακάτω.

Ορισμός ((d,t)-delineable). Λέμε ότι μία κλάση μηχανισμών $\mathcal{C}$ είναι (d,t)-delineable εάν:

1. Η κλάση $\mathcal{C}$ αποτελείται από μηχανισμούς που παραμετροποιούνται από διανύσματα $\boldsymbol{p}$ από ένα σύνολο $\mathcal{P} \subseteq \mathbb{R}^d$ και
2. Για κάθε $\boldsymbol{v}$ στο πεδίο τιμών της κατανομής D των συναρτήσεων αξιών των παικτών, υπάρχει ένα σύνολο $\mathcal{H}$ το t υπερεπιπέδων τέτοιων ώστε για κάθε συνεκτική περιοχή $\mathcal{P}'$ του $\mathcal{P} \setminus \mathcal{H}$ το εισόδημα του μηχανισμού στο $\boldsymbol{v}$ να είναι γραμμική συνάρτηση των παραμέτρων του μηχανισμού.

Διαισθητικά, όσο μεγαλύτερος είναι ο αριθμός t των υπερεπιπέδων που χωρίζουν μία κλάση μηχανισμών σε γραμμικές περιοχές, τόσο μεγαλύτερη είναι η εκφραστική ικανότητα της οικογένειας αυτής, και κατ' επέκταση τόσο μεγαλύτερο είναι το generalization error της κλάσης αυτής. Οι συγγραφείς της εργασίας αυτής θεμελίωσαν τη διαίσθηση αυτή με το παρακάτω θεώρημα.

Θεώρημα. Εάν μία κλάση μηχανισμών είναι (d, t)-delineable, η ψευδο-διάστασή της είναι $O(d \log(dt))$.

Για την Unit-Bundling κλάση που εισάγαμε αποδείξαμε το παρακάτω θεώρημα:

Θεώρημα 6.1 Η κλάση των unit bundling δημοπρασιών αναβαλλόμενης αποδοχής για n παίκτες και m αντίτυπα του αγαθού διαθέσιμα προς πώληση είναι $(2n, nm)$-delineable.

Η απόδειξη αυτή είχε αρκετές τεχνικές λεπτομέρειες, ωστόσο η λογική της ήταν σχετικά απλή: Για οποιοδήποτε διάνυσμα $\boldsymbol{v}$, το εισόδημα του μηχανισμού μπορεί να εκφραστεί ως το άθροισμα των πληρωμών των n παικτών. Έπειτα, μπορεί κανείς να δείξει ότι για κάθε παίκτη, για κάθε ένα από τα $m+1$ πιθανά bundle sizes για τον παίκτη αυτόν, υπάρχουν μόνο 2 γραμμικές περιοχές: Η περιοχή στην οποία αγοράζει το bundle, και η περιοχή στην οποία δεν το αγοράζει, οι οποίες διαχωρίζονται από το υπερεπίπεδο στο οποίο η τιμή του bundle είναι ακριβώς ίση με την αξία του παίκτη για αυτό. Πολλαπλασιάζοντας με $m+1$ για όλα τα πιθανά bundle sizes, υπάρχουν ακριβώς $m+1$ τέτοια υπερεπίπεδα για κάθε παίκτη. Τέλος, πολλαπλασιάζοντας με το συνολικό αριθμό των παικτών, προκύπτει ότι για κάθε $\boldsymbol{v}$ υπάρχουν το πολύ $n(m+1)$ υπερεπίπεδα που χωρίζουν το χώρο $\mathcal{P}$ των παραμέτρων του μηχανισμού σε περιοχές όπου το εισόδημα είναι γραμμικό.

Από τα δύο παραπάνω θεωρήματα άμεσα προκύπτει:

Θεώρημα 6.2 Η ψευδο-διάσταση της κλάσης των unit bundling δημοπρασιών αναβαλλόμενης αποδοχής για n παίκτες και m αντίτυπα του αγαθού διαθέσιμα προς πώληση είναι $O(n \log(nm))$.

Το να θεμελιώσουμε ένα representation error bound για αυτή την οικογένεια ήταν πιο απαιτητικό. Η βασική πρόκληση, όπως αναφέραμε και στην εισαγωγή, είναι ότι τώρα πρέπει να συγκρίνουμε το αναμενόμενο εισόδημα του βέλτιστου μηχανισμού από την κλάση που προτείνουμε με την αναμενόμενη κοινωνική ωφέλεια του VCG. Για να το επιτύχουμε αυτό χρησιμοποιήσαμε μία αρκετά εξωτική τεχνική: Εισάγαμε έναν νέο μηχανισμό, τον "εκ των προτέρων βέλτιστο μηχανισμό", την αναμενόμενη κοινωνική ωφέλεια του οποίου χρησιμοποιήσαμε ως διεπαφή ανάμεσα στο αναμενόμενο εισόδημα του βέλτιστου μηχανισμού της δικής μας κλάσης και την αναμενόμενη κοινωνική ωφέλεια του VCG. Στη συνέχεια φράξαμε τις 2 αυτές διαφορές, το άθροισμα των οποίων αποτελεί απευθείας ένα άνω φράγμα στο representation error της κλάσης μας. Αρχικά ορίζουμε τον μηχανισμό αυτόν:

Ορισμός (εκ τον προτέρων βέλτιστος μηχανισμός - $\mathcal{A}$). Ο $\mathcal{A}$ είναι ένας υποθετικός μηχανισμός. Διαθέτει τέλεια γνώση των κατανομών όλων των παικτών και επιλέγει, χωρίς να λάβει καθόλου υπόψιν του τις πραγματικές αξίες των παικτών, την σταθερή ανάθεση των αντικειμένων σε παίκτες με τη μέγιστη αναμενόμενη κοινωνική αξία.

Στη συνέχεια, αποδείξαμε χρησιμοποιώντας ένα αρκετά τεχνικό πόρισμα της εργασίας "Revenue Optimization with Approximate Bid Predictions" ότι για το αναμενόμενο εισόδημα του βέλτιστου μηχανισμού από την κλάση $\mathcal{UB}$ ισχύει:

Θεώρημα 6.3 Για κάθε κατανομή F των συναρτήσεων αξιών των παικτών, η κλάση των unit bundling συναρτήσεων περιέχει ένα μηχανισμό $\mathcal{M}$ για το αναμενόμενο εισόδημα του οποίου ισχύει:

$$\mathbb{E}_{\boldsymbol{v}}[\mathrm{Rev}(\mathcal{M},\boldsymbol{v})] \geq \mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathcal{A},\boldsymbol{v})] - \sum_{i=1}^{n}(3\widehat{B}_i)^{1/3}\widehat{\sigma_i}^{2/3} \tag{1.3}$$

όπου $\widehat{B_i}$: Η αναμενόμενη αξία του παίκτη i για τον αριθμό αντιτύπων του αγαθού που του ανατίθενται από το μηχανισμό $\mathcal{A}$,
$\widehat{\sigma_i}^2$: Η διασπορά της συσσωρευτικής κατανομής της αξίας του παίκτη i για τον αριθμό αντικειμένων που του ανατίθενται από τον μηχανισμό $\mathcal{A}$, δηλαδή $\widehat{\sigma_i}^2 = \sum_{j=1}^{s_i^*}\sigma_{i,j}^2$.

Η απόδειξη μπορεί να συνοψιστεί ως εξής: Η κλάση $\mathcal{UB}$ περιέχει έναν μηχανισμό $\mathcal{M}$ που επιλέγει για κάθε παίκτη το ίδιο μέγεθος bundle με την ανάθεση του παίκτη αυτού στο μηχανισμό $\mathcal{A}$ και στη συνέχεια θέτει, για κάθε παίκτη, τη βέλτιστη τιμή για το bundle αυτό. Έπειτα το μόνο που μένει είναι να φραχτεί για κάθε παίκτη η διαφορά ανάμεσα στην αξία του για το bundle αυτό και της αναμενόμενης πληρωμής του, δεδομένου ότι ο μηχανισμός έχει επιλέξει τη βέλτιστη τιμή για το bundle αυτό. Υπάρχει όμως ένα πόρισμα της εργασίας που προαναφέραμε που φράζει, για κάθε κατανομή, ακριβώς αυτή τη διαφορά.
Τώρα το μόνο που μένει για να εδραιώσουμε ένα representation error bound για την $\mathcal{UB}$ κλάση είναι να φράξουμε τη διαφορά ανάμεσα στην αναμενόμενη κοινωνική ωφέλεια του VCG και εκείνη του $\mathcal{A}$. Με το παρακάτω θεώρημα δείξαμε ακριβώς αυτό για Gaussian και Sub-Gaussian κατανομές.

Θεώρημα 6.5 Στο περιβάλλον της υποενότητας αυτής, για την αναμενόμενη κοινωνική ωφέλεια του εκ των προτέρων βέλτιστου μηχανισμού $\mathcal{A}$ ισχύει:

$$\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathcal{A},\boldsymbol{v})] \geq \mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathrm{VCG},\boldsymbol{v})] - \sum_{i=1}^{m}\overline{\sigma}_i\sqrt{2\log(nm)}, \tag{1.4}$$

όπου $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_{mn}$ είναι ένα total ordering των τυπικών αποκλίσεων των κατανομών όλων των nm οριακών αξιών των παικτών.

Η κεντρική ιδέα της απόδειξης ήταν ότι για κάθε διάνυσμα αξιών $\boldsymbol{v}$, η ανάθεση του $\mathcal{A}$ μπορεί να 'μετατραπεί' στην ανάθεση του VCG με το πολύ m 'μετακινήσεις' αντιτύπων, όπου με μία 'μετάθεση' αφαιρούμε ένα αντικείμενο από έναν παίκτη που έλαβε περισσότερα στον $\mathcal{A}$ από ότι στον VCG και το αναθέτουμε σε κάποιον που έλαβε λιγότερα. Η αναμενόμενη διαφορά των 2 μηχανισμών είναι ακριβώς το αναμενόμενο 'κέρδος' σε κοινωνική ωφέλεια του VCG από τις μετακινήσεις αυτές. Τέλος για Gaussian και Sub-Gaussian κατανομές φράξαμε το αναμενόμενο κέρδος των μετακινήσεων αυτών, που αμέσως συνεπάγεται ότι φράξαμε την αναμενόμενη διαφορά της κοινωνικής ωφέλειας των δύο μηχανισμών.
Συνδυάζοντάς τα θεωρήματα 6.2, 6.3 και 6.5 άμεσα προκύπτει το ζητούμενο πόρισμα:

Πόρισμα 6.3 Στο περιβάλλον της υποενότητας αυτής, για τον μηχανισμό $\mathcal{M} \in \mathcal{UB}$ με το εμπειρικά μέγιστο εισόδημα σε ένα σύνολο N δειγμάτων των κατανομών των παικτών, ισχύει:

$$\mathbb{E}_{\boldsymbol{v}}[\mathrm{Rev}(\mathcal{M}, \boldsymbol{v})] \geq \mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathrm{VCG}, \boldsymbol{v})] - \sum_{i=1}^{n}(3\widehat{B}_i)^{1/3}\widehat{\sigma}_i{}^{2/3} - \sum_{i=1}^{m}\overline{\sigma}_i\sqrt{2\log(nm)}$$
$$- O\left(U\sqrt{n\log(nm)/N} + U\sqrt{\ln(1/\delta)/N}\right) \quad (1.5)$$

Αξίζει να σχολιάσουμε το αποτέλεσμα αυτό: Για παίκτες με submodular συναρτήσεις αξίας, κανένας weakly group strategyproof μηχανισμός δε μπορεί να εγγυηθεί σε κάθε διάνυσμα αξιών $\boldsymbol{v}$ προσέγγιση της κοινωνικής ωφέλειας του VCG καλύτερη από $\frac{1}{\sqrt{2}}$. Εμείς όμως ασχοληθήκαμε με το ακόμα δυσκολότερο πρόβλημα του να συγκρίνουμε το εισόδημα του μηχανισμού που προτείνουμε απευθείας με την κοινωνική ωφέλεια του VCG. Δείξαμε ότι εάν οι αξίες των παικτών ακολουθούν Gaussian και Sub-Gaussian κατανομές, δοθέντων κάποιων δειγμάτων από τις κατανομές αυτές, είναι δυνατό να μάθει κανείς μία δημοπρασία αναβαλλόμενης αποδοχής με αναμενόμενο εισόδημα ίσο με την αναμενόμενη κοινωνική ωφέλεια του VCG μείον 3 αθροιστικούς όρους:

1. Τον όρο $\sum_{i=1}^{n}(3\widehat{B}_i)^{1/3}\widehat{\sigma}_i{}^{2/3}$: Εάν δεν υπήρχε η τρίτη ρίζα πάνω από κάθε όρο του αθροίσματος και οι συνολικές τυπικές αποκλίσεις όλων των παικτών για τον αριθμό αντιτύπων που τους δίνονται από τον $\mathcal{A}$ ήταν μονάδα, τότε αυτό το άθροισμα θα ήταν ίσο με 3 φορές την αναμενόμενη κοινωνική ωφέλεια του VCG. Βεβαίως, για 'εύλογες' τιμές παραμέτρων των κατανομών αξιών των παικτών, η τρίτη ρίζα καθιστά το άθροισμα αυτό τάξεις μεγέθους μικρότερο από την αναμενόμενη κοινωνική ωφέλεια του VCG.
2. Τον όρο $\sum_{i=1}^{m}\overline{\sigma}_i\sqrt{2\log(nm)}$: Ο όρος αυτός εξαρτάται γραμμικά από τις τυπικές αποκλίσεις των κατανομών αξιών των παικτών, άρα υπογραμμικά από τις διασπορές τους, υπογραμμικά από τον αριθμό των παικτών, και 'σχεδόν' γραμμικά από τον αριθμό των αντικειμένων διαθέσιμων προς πώληση. Υπό την εύλογη υπόθεση ότι για κάθε κατανομή αξίας των παικτών, η αναμενόμενη τιμή της κατανομής αυτής είναι σημαντικά μεγαλύτερη από την τυπική απόκλιση της κατανομής αυτής, και αυτός ο όρος είναι τάξη μεγέθους μικρότερος από την αναμενόμενη κοινωνική ωφέλεια του VCG.
3. Τον όρο $O\left(U\sqrt{n\log(nm)/N} + U\sqrt{\ln(1/\delta)/N}\right)$: Το υπόλοιπό 'κομμάτι' αυτού του όρου πέρα από το $n\log(nm)$ προέρχεται απευθείας από το θεώρημα ομοιόμορφης σύγκλισης του Pollard, και είναι κοινό σε σχεδόν κάθε προσέγγιση ανάλυσης δειγματικής πολυπλοκότητας. Για να αιτιολογήσουμε λοιπόν για το εάν αυτός ο όρος είναι καλός, αρκεί να επικεντρωθούμε στο $n\log(nm)$, την ψευδο-διάσταση δηλαδή που δείξαμε για την προτεινόμενη κλάση $\mathcal{UB}$. Υπό μη-ανώνυμες τιμές για τα bundles, δηλαδή διαφορετικές τιμές για κάθε παίκτη, η εξάρτηση της ψευδο-διάστασης από τον αριθμό των παικτών n είναι αναμενόμενη. Τέλος, για συναρτήσεις αξιών των παικτών που εξαρτώνται άμεσα από τον αριθμό των αντικειμένων που λαμβάνουν, αναμενόμενο είναι η δειγματική πολυπλοκότητα του μηχανισμού να εξαρτάται και από τον αριθμό των αντικειμένων m. Αυτός ο όρος είναι λοιπόν, κατά μία έννοια, 'όσο καλός' θα μπορούσε να είναι. Για μία περισσότερο εις βάθος ανάλυση αυτών των ισχυρισμών προτείνουμε στον ενδιαφερόμενο αναγνώστη την εργασία "A General Theory of Sample Complexity for Multi-Item Profit Maximization".

1.4 Περιορίζοντας τον Αριθμό Δειγμάτων

Τα αποτελέσματά μας που παρουσιάστηκαν στην προηγούμενη υποενότητα ήταν πολύ υποσχόμενα. Για κάθε περιβάλλον που μελετήσαμε, δείξαμε ότι δοθέντων αρκετών δειγμάτων, είναι εφικτό να μάθει κανείς μία δημοπρασία αναβαλλόμενης υποδοχής με πολύ υψηλό αναμενόμενο εισόδημα. Σε κάποιες αγορές μεγάλης σημασίας, όπως για παράδειγμα οι στις αγορές απονομής φάσματος όπου χρησιμοποιούνται δημοπρασίες αναβαλλόμενης αποδοχής, οι δημοπρασίες μπορεί να είναι αρκετά σπάνιες. Για τέτοια προβλήματα, η υπόθεση ότι διαθέτει κανείς μη-σταθερό αριθμό δειγμάτων, ή άλλες γνώσεις για τις κατανομές αξιών των παικτών, ίσως να είναι μη ρεαλιστική. Ένα λογικό επόμενο βήμα είναι να μελετήσουμε πώς μπορούμε να κατασκευάσουμε δημοπρασίες αναβαλλόμενης αποδοχής που, χρησιμοποιώντας έναν πολύ μικρό, σταθερό αριθμό δειγμάτων και καμία άλλη γνώση για τις κατανομές των παικτών, πετυχαίνουν αναμενόμενο εισόδημα που είναι μία σταθερή προσέγγιση αυτού που δείξαμε ότι μπορεί να επιτευχθεί με περισσότερα δείγματα. Και πάλι θα ασχοληθούμε τόσο με single- όσο και multi-parameter περιβάλλοντα, όμοια σχεδόν με αυτά της προηγούμενης ενότητας.

1.4.1 Single-Parameter Περιβάλλοντα

Το περιβάλλον που θα μελετήσουμε σε αυτή την υποενότητα είναι το ίδιο με την αντίστοιχη περίπτωση όπου είχαμε πολλά δείγματα. Η μόνη διαφορά είναι ότι τώρα δεν απαιτούμε οι κατανομές αξιών των παικτών να είναι περιορισμένες σε κάποιο διάστημα $[1, H]$, αλλά μόνο ότι είναι κανονικές. Αυτό είναι μία σημαντικά ασθενέστερη υπόθεση: Κάθε φραγμένη κατανομή είναι και κανονική.

Το Περιβάλλον (Παίκτες με αθροιστικές αξίες, κανονικές κατανομές)

- n παίκτες.
- Υπάρχουν m αντίτυπα του αγαθού διαθέσιμα προς πώληση.
- Κάθε παίκτης i έχει μία γνωστή εκ των προτέρων στο μηχανισμό *ζήτηση* d_i.
- Κάθε παίκτης i έχει αθροιστική αξία, μέχρι τη ζήτησή του d_i:
  Για κάθε παίκτη i, η αξία του για x_i αντίτυπα του αγαθού, δεδομένου ότι $x_i \leq d_i$, είναι $x_i \cdot v_i$.
- Η αξία του κάθε παίκτη i ανά αντίτυπο του αγαθού, v_i, ακολουθεί κάποια κατανομή με συνάρτηση πυκνότητας πιθανότητας $f(\cdot)$ και virtual valuation function $\phi_i(\cdot)$.
- Οι κατανομές αυτές είναι κανονικές.

Η προσέγγισή μας σε αυτό το περιβάλλον ήταν εμπνευσμένη από την εργασία των Hartline και Roughgarden με τίτλο "Simple versus Optimal Mechanisms". Στη δική μας περίπτωση ωστόσο, απαιτούνταν κάποιες μη-τετριμμένες τροποποιήσεις. Οι Hartline και Roughgarden στην εργασία τους αυτή εισήγαγαν την έννοια του 'διπλούτυπου' (duplicated) περιβάλλοντος.

Ορισμός (Διπλότυπο Περιβάλλον). Κάθε παίκτης i με κατανομή F_i αντικαθίσταται από ένα ζεύγος παικτών i, i' των οποίων οι αξίες είναι i.i.d. δείγματα από την κατανομή F_i. Οι εφικτές αναθέσεις στο διπλότυπο περιβάλλον είναι αυτές που ικανοποιούν:

1. Από κάθε ζευγάρι i, i' το πολύ ένας παίκτης έχει μη-μηδενική ανάθεση.
2. Η ανάθεση, αν ερμηνευθεί με τον φυσικό της τρόπο ως η ανάθεση στο αρχικό περιβάλλον, είναι εφικτή στο περιβάλλον αυτό.

Μία άλλη σημαντική έννοια που εισήγαγαν οι Hartline και Roughgarden ήταν αυτή των "ανάλογων μηχανισμών" (commensurate mechanisms). Οι Hartline και Roughgarden την εισήγαγαν για δυαδικά περιβάλλοντα, ενώ εμείς την επεκτείναμε με τον πιο φυσικό τρόπο σε αυθαίρετα single-parameter περιβάλλοντα. Παρακάτω παρατίθεται απευθείας ο επεκταμένος ορισμός.

Ορισμός (Single-Parameter Ανάλογος Μηχανισμός). Έστω $\mathcal{M}$ και $\mathcal{M}'$ δύο μηχανισμοί για κάποιο single-parameter περιβάλλον. Έστω $x_i(\boldsymbol{v}), x'_i(\boldsymbol{v})$ η ανάθεση στον παίκτη i των 2 μηχανισμών στο προφίλ αξιών $\boldsymbol{v}$. Ο μηχανισμός $\mathcal{M}$ είναι single-parameter ανάλογος με το μηχανισμό $\mathcal{M}'$ εάν:

$$\mathbb{E}_{\boldsymbol{v}}\left[\sum_{i:x_i(\boldsymbol{v})\neq x'_i(\boldsymbol{v})} x_i(\boldsymbol{v})\phi_i(v_i)\right] \geq 0 \tag{C1}$$

και

$$\mathbb{E}_{\boldsymbol{v}}\left[\sum_{i\in\mathcal{N}} p_i(\boldsymbol{v})\right] \geq \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i:x_i(\boldsymbol{v})\neq x'_i(\boldsymbol{v})} x'_i(\boldsymbol{v})\phi_i(v_i)\right], \tag{C2}$$

όπου $p_i(\boldsymbol{v})$: Η πληρωμή του παίκτη i στον $\mathcal{M}$ για το προφίλ αξιών $\boldsymbol{v}$, για ανάθεση $x_i(\boldsymbol{v})$.
Ο λόγος που εισάγαμε την έννοια των single-parameter ανάλογων μηχανισμών είναι ότι μπορούμε, αντίστοιχα την πρωτότυπη εργασία, να αποδείξουμε ότι ισχύει το παρακάτω λήμμα.

Λήμμα 7.1 Εάν ένας μηχανισμός $\mathcal{M}$ είναι single-parameter ανάλογος με έναν μηχανισμό $\mathcal{M}'$, τότε για το αναμενόμενο εισόδημα των δύο μηχανισμών ισχύει:

$$\mathbb{E}_{\boldsymbol{v}}\left[\mathrm{Rev}(\mathcal{M}, \boldsymbol{v})\right] \geq \frac{1}{2}\cdot\mathbb{E}_{\boldsymbol{v}}\left[\mathrm{Rev}(\mathcal{M}', \boldsymbol{v})\right]. \tag{1.6}$$

Η απόδειξη στηρίζεται στο ότι κάθε μία από τις συνθήκες C1 και C2 αρκούν για να δείξουν ότι το αναμενόμενο εισόδημα του μηχανισμού $\mathcal{M}'$ από τους παίκτες που ταυτίζονται ή δεν ταυτίζονται αντίστοιχα οι αναθέσεις των δύο μηχανισμών είναι μικρότερο ή ίσο του αναμενόμενου συνολικού εισοδήματος του μηχανισμού $\mathcal{M}$.
Ο μόνος λόγος που αναφέραμε την έννοια του διπλοτύπου περιβάλλοντος σε αυτή την διπλωματική εργασία είναι ότι μπορέσαμε να δείξουμε ότι ισχύει το εξής:

Λήμμα 7.2 Στο περιβάλλον μίας δημοπρασίας πολλαπλών αντιτύπων όπου οι παίκτες έχουν αθροιστικές αξίες που ακολουθούν κάποια κανονική κατανομή, ο VCG μηχανισμός στο διπλότυπο περιβάλλον είναι single-parameter ανάλογος με τον βέλτιστο μηχανισμό, Myerson, στο αρχικό περιβάλλον.

Η κεντρική ιδέα τώρα είναι η εξής: Σε τέτοιου είδους περιβάλλοντα, οι Γκατζέλης, Μαρκάκης και Roughgarden έδειξαν ότι ο VCG μηχανισμός μπορεί να υλοποιηθεί ως δημοπρασία αναβαλλόμενης αποδοχής. Εάν λοιπόν ο δημοπράτης μπορούσε με κάποιο τρόπο να 'διπλασιάσει' τους παίκτες, και να εφαρμόσει σε αυτό το περιβάλλον την υλοποίηση του VCG ως δημοπρασία αναβαλλόμενης αποδοχής, τότε θα είχε αναμενόμενο εισόδημα το μισό από ότι με το μηχανισμό Myerson, δηλαδή το μισό από το βέλτιστο. Ο δημοπράτης προφανώς δε μπορεί να διπλασιάσει τους παίκτες. Αυτό όμως που μπορεί να κάνει είναι, εάν έχει ένα δείγμα από την αξία του κάθε παίκτη, να εφαρμόσει τον μηχανισμό αυτό αντιμετωπίζοντας το κάθε δείγμα v_i ως το διπλότυπο i' του αντίστοιχου παίκτη i, προσομοιώνοντας με αυτό τον τρόπο τους διπλότυπους παίκτες. Από κάθε ζεύγος i, i', όποτε κερδίζει ο αυθεντικός παίκτης, αυτός πληρώνει κανονικά το αντίστοιχο ποσό. Από την άλλη μεριά, όποτε κερδίζει το δείγμα που αντιμετωπίζουμε ως το διπλότυπό του, ο δημοπράτης δεν πληρώνεται, αφού στην πραγματικότητα δεν υπάρχει ο αντίστοιχος παίκτης αλλά απλώς ένα δείγμα της αξίας του. Έτσι, λόγω συμμετρίας, από κάθε ζεύγος παικτών, σε σχέση με το εάν υπήρχαν οι διπλότυποι παίκτες, με πιθανότητα $\frac{1}{2}$ ο μηχανισμός που προσομοιώνει διπλότυπους παίκτες δεν πληρώνεται, επομένως έχει αναμενόμενο εισόδημα ακριβώς το μισό από ότι ο VCG στο (πραγματικό) διπλοτυπο περιβάλλον. Συνολικά λοιπόν, ο μηχανισμός αυτός με ένα δείγμα πετυχαίνει αναμενόμενο εισόδημα τουλάχιστον $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ του βέλτιστου.

Πόρισμα 7.1 Σε κάθε δημοπρασία πολλαπλών αντιτύπων όπου οι παίκτες έχουν αθροιστικές αξίες που ακολουθούν κανονικές κατανομές, υπάρχει μία γενικευμένη δημοπρασία αναβαλλόμενης αποδοχής η οποία, χρησιμοποιώντας ένα δείγμα των αξιών των παικτών, επιτυγχάνει αναμενόμενο εισόδημα που είναι τουλάχιστον ένα τέταρτο το βέλτιστου.

1.4.2 Multi-Parameter Περιβάλλοντα

Το περιβάλλον που θα μελετήσουμε σε αυτή την υποενότητα είναι ακριβώς το ίδιο με την αντίστοιχη περίπτωση όπου είχαμε πολλά δείγματα. Όλοι οι παίκτες έχουν submodular συναρτήσεις αξίας, οι οποίες απορρέουν από Gaussian και Sub-Gaussian κατανομές. Συγκεκριμένα:

Το Περιβάλλον (Παίκτες με submodular αξίες)

- n παίκτες.
- Υπάρχουν m αντίτυπα του αγαθού διαθέσιμα προς πώληση.
- Κάθε παίκτης i δηλώνει τις m *οριακές του αξίες* στο μηχανισμό, $v_{i,1}, v_{i,2}, \ldots, v_{i,m}$, όπου $v_{i,j}$: Η αξία που έχει για τον παίκτη i να αποκτήσει το j-οστό αντίτυπό του του αγαθού, δεδομένου ότι έχει εξασφαλίσει ήδη $j-1$ αντίτυπα. Για κάθε παίκτη i, η αξία του για x_i αντίτυπα του αγαθού, δεδομένου ότι $x_i \leq d_i$, είναι $x_i \cdot v_i$.
- Η συνάρτηση αξίας κάθε παίκτη είναι submodular: $v_{i,1} \geq v_{i,2} \geq \cdots \geq v_{i,m}$ $\forall i$
- Κάθε οριακή αξία $v_{i,j}$ ακολουθεί κάποια Gaussian ή Sub-Gaussian κατανομή $F_{i,j}$.

Για το περιβάλλον αυτό προτείνουμε έναν σχετικά απλό μηχανισμό, ο οποίος σε αντίθεση με την προηγούμενη υποενότητα, χρησιμοποιεί τώρα 2 δείγματα. Για το λόγο αυτό τον ονομάσαμε το μηχανισμό 2 δειγμάτων (Two Samples Mechanism - $\mathcal{TSM}$):

Algorithm 1: Ο Μηχανισμός Δύο Δειγμάτων ($\mathcal{TSM}$)

Συνέλεξε 2 δείγματα από τις κατανομές όλων των παικτών;
Εφάρμοσε τον VCG στο πρώτο δείγμα. Έστω $\boldsymbol{x} = (x_1, x_2, \ldots x_n)$ η ανάθεση της εκτέλεσης αυτής;
Για κάθε παίκτη i προσδιόρισε την αξία του r_i για το bundle x_i αντιτύπων του αγαθού στο δεύτερο δείγμα;
Στη δημοπρασία προσέφερε σε κάθε παίκτη i ένα bundle x_i αντιτύπων σε τιμή $0.85 \cdot r_i$;

Η προσέγγισή μας στην ανάλυση του μηχανισμού αυτού είναι ανάλογη με την περίπτωση όπου διαθέταμε περισσότερα δείγματα για το ίδιο περιβάλλον. Αρχικά θα φράξουμε τη διαφορά μεταξύ της κοινωνικής ωφέλειας του μηχανισμού αυτού και της αναμενόμενης κοινωνικής ωφέλειας του VCG. Έπειτα θα φράξουμε τη διαφορά μεταξύ της αναμενόμενης κοινωνικής ωφέλειας και του αναμενόμενου εισοδήματος του μηχανισμού αυτού. Το άθροισμα των δύο αυτών διαφορών απαιτεί ένα έμμεσο φράγμα για τη διαφορά ανάμεσα στην αναμενόμενη κοινωνική ωφέλεια του VCG και το αναμενόμενο εισόδημα του μηχανισμού που προτείνουμε.

Λήμμα 7.1 Στο περιβάλλον αυτής της υποενότητας, για την αναμενόμενη κοινωνική ωφέλεια της αναμενόμενης ανάθεσης $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ που ο $\mathcal{TSM}$ μηχανισμός καθόρισε στο πρώτο δείγμα ισχύει:

$$\mathbb{E}_{\boldsymbol{v}}\left[\mathrm{SW}(\boldsymbol{x}, \boldsymbol{v})\right] = \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i=1}^{n}\sum_{j=1}^{x_i} v_{i,j}\right] \geq \mathbb{E}_{\boldsymbol{v}}\left[\mathrm{SW}(\mathcal{A}, \boldsymbol{v})\right] - 2\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i, \tag{1.7}$$

όπου $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_{mn}$ είναι ένα total ordering των τυπικών αποκλίσεων των κατανομών όλων των nm οριακών αξιών των παικτών.

Η απόδειξη αυτή είναι παρόμοια με εκείνη του θεωρήματος 6.5. Για Guassian και Sub-Gaussian κατανομές, μπορούμε να δείξουμε ότι το αναμενόμενο μέγιστο που m από τις nm οριακές τιμές μπορεί να απέχουν από την αναμενόμενη τους τιμή είναι $\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i$. Αυτό όμως ισχύει και για την κατανομή $\boldsymbol{x}$ που επέλεξε ο VCG στο πρώτο δείγμα. Η αναμονή είναι ότι η ανάθεση αυτή στο πρώτο δείγμα να είχε το πολύ μέχρι και $\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i$ μεγαλύτερη συνολική αξία από την αναμενόμενη της. Αντίστοιχα, η αναμονή για την ανάθεση του εκ των προτέρων βέλτιστου μηχανισμού είναι ότι στο πρώτο δείγμα μπορεί να ήταν το πολύ $\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i$ χαμηλότερη από την αναμενόμενη τιμή της. Αυτά τα δύο, σε συνδυασμό με το γεγονός ότι στο πρώτο δείγμα η ανάθεση $\boldsymbol{x}$ έχει μεγαλύτερη ή ίση κοινωνική ωφέλεια από οποιαδήποτε άλλη ανάθεση αφού είναι αυτή που επιλέγει ο VCG αμέσως αποδεικνύουν το ζητούμενο.

Αυτό το λήμμα, σε συνδυασμό με το φράγμα ανάμεσα στην αναμενόμενη κοινωνική ωφέλεια του $\mathcal{A}$ και του VCG αρκούν ώστε να εδραιώσουμε ένα κάτω φράγμα για την αναμενόμενη κοινωνική ωφέλεια του μηχανισμού $\mathcal{TSM}$ που προτείνουμε. Για να ολοκληρωθεί η ανάλυσή μας, αρκεί να φράξουμε τη διαφορά ανάμεσα στην αναμενόμενη κοινωνική ωφέλεια του $\mathcal{TSM}$ και στο αναμενόμενο εισόδημά του. Αυτό επιτυγχάνεται με το τελευταίο μας λήμμα:

Λήμμα 7.5 Έστω $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ η ανάθεση που ο $\mathcal{TSM}$ μηχανισμός καθόρισε στο πρώτο δείγμα, και έστω $\widetilde{B}_i, \widetilde{\sigma}_i^{\,2}$ η αναμενόμενη τιμή και η διασπορά της $\widetilde{F}_i$, της αναμενόμενης αξίας του παίκτη i για x_i αντίτυπα του αγαθού. Τότε, εάν οι κατανομές αξιών όλων των παικτών είναι monotone hazard rate, για το αναμενόμενο εισόδημα του $\mathcal{TSM}$ ισχύει:

$$\mathbb{E}_{\boldsymbol{v}}[\mathrm{Rev}(\mathcal{TSM}, \boldsymbol{v})] \geq 0.589 \cdot \mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\boldsymbol{x}, \boldsymbol{v})] - 0.589 \cdot \sum_{i=1}^{n}(3\widetilde{B}_i)^{1/3}\widetilde{\sigma}_i^{\,2/3} \tag{1.8}$$

Η απόδειξη του λήμματος είναι αρκετά τεχνική και παραλείπεται από την περίληψη αυτή. Συνδυάζοντας τα λήμματα 7.1 και 7.5 με το φράγμα ανάμεσα στην αναμενόμενη κοινωνική ωφέλεια του $\mathcal{A}$ και του VCG που δείξαμε στο θεώρημα 6.5 προκύπτει το τελευταίο μας πόρισμα:

Πόρισμα 7.3 Στο περιβάλλον μίας δημοπρασίας πολλαπλών αντιτύπων όπου οι παίκτες έχουν submodular αξίες που ακολουθούν Gaussian και Sub-Gaussian αξίες, για το αναμενόμενο εισόδημα του μηχανισμού $\mathcal{TSM}$, όπου η αναμονή είναι πάνω στα δύο δείγματα που ο μηχανισμός χρησιμοποίησε, ισχύει:

$$\mathbb{E}_{\boldsymbol{v}}[\mathrm{Rev}(\mathcal{TSM}, \boldsymbol{v})] \geq 0.589 \cdot \mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathrm{VCG}, \boldsymbol{v})] - 1.767\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i - 0.589 \cdot \sum_{i=1}^{n}(3\widetilde{B}_i)^{1/3}\widetilde{\sigma}_i^{\,2/3} \tag{1.9}$$

Αξίζει να σχολιάσουμε πόσο καλό είναι το αποτέλεσμα αυτό: Για submodular bidders, κανένας weakly group-strategyproof μηχανισμός δε μπορεί να εγγυηθεί ότι η κοινωνική του ωφέλεια θα είναι πάνω από $1/\sqrt{2}$ προσέγγιση της βέλτιστης κοινωνικής ωφέλειας, δηλαδή αυτής του VCG. Εμείς, προτείναμε έναν μηχανισμό ο οποίος, με εύλογες-φυσικές υποθέσεις για τις κατανομές αξιών των παικτών, χρησιμοποιώντας μόνο δύο δείγματα από τις κατανομές τους πετυχαίνει, on expectation πάνω στα δύο δείγματα που χρησιμοποίησε, αναμενόμενο εισόδημα που είναι 0.589-προσέγγιση της αναμενόμενης κοινωνικής ωφέλειας του VCG, μείον κάτι αθροιστικούς όρους που όπως εξηγήσαμε, για εύλογες κατανομές είναι τάξη μεγέθους μικρότεροι από το αναμενόμενο εισόδημα.

Μία ενδιαφέρουσα τελευταία παρατήρηση είναι ότι στην περίπτωση που έχουμε πολύ περιορισμένα δείγματα, τόσο στο single- οσο και στο multi-parameter περιβάλλον, ουσιαστικά χρησιμοποιούμε τα δείγματα αυτά ως μία προσέγγιση των κατάλληλων παραμέτρων για ένα μηχανισμό από την αντίστοιχη οικογένεια που προτείναμε για το ίδιο περιβάλλον στην περίπτωση όπου είχαμε πολλά δείγματα. Συγκεκριμένα, στο single-parameter περιβάλλον, χρησιμοποιήσαμε ουσιαστικά ένα δείγμα από την αξία του κάθε παίκτη i ως μία προσέγγιση του βέλτιστου κατωφλίου του $l_{i,0}$ σε μία δημοπρασία ενός επιπέδου. Στο multi-parameter περιβάλλον από την άλλη μεριά, χρησιμοποιήσαμε ένα δείγμα για να προσδιορίσουμε μία προσέγγιση των βέλτιστων bundle sizes ενός unit bundling μηχανισμού, και ένα δεύτερο δείγμα για να προσδιορίσουμε τις βέλτιστες τιμές για τα bundles αυτά.

Chapter 2

Introduction

Imagine you own a traveling agency. The summer has arrived and you want to hire buses in order to transport people from the city to the beach and back. As a private company, your aim is to make as much profit from this endeavor as possible. How should you price your seats? Set a price too low and all the buses will be filled, but you won't make too much money. Set a price too high and there will be empty seats, which imply lost revenue. After some thought, the logical next step is to try auctioning-off the seats. The thought process is that you have a finite number of seats, and the potential clients will bid for them. Those with the higher bids will get a seat, at the price they bid, until the buses are filled. At first glance this seems like a good idea: Now the buses will be filled and the prices will be high due to the competition between the travelers for the seats. But after some more careful examination, you realize your plan has one fatal flaw: The travelers will realize that if they bid too high they will get a seat, but someone else may have gotten a seat for a much lower price. Compared to the passenger that bid less and still got a seat, they overpaid. After this realization, the potential buyers will start bidding less than the intended to in an attempt to save money, driving down the prices.

As if this problem was not hard enough to solve, imagine that instead of buses it involved airplanes. Once again, your aim is to maximize your profit for the seats sold. Now there are different types of seats (economy class, extended legroom, business class) that correspond to different levels of service. Additionally, the potential travelers can bid for more than one seat, as they may want to go on a family trip. But families are special: Either all members of the family will travel, or none of them, and if you are to service them then they need to be assigned adjacent seats. Finally, people at business class paid a higher price, for a higher quality of service. They don't want to be woken up to the sound of crying babies.

It becomes apparent that in such complex and unprincipled systems the people's preferences may collude, in the sense that they cannot all be satisfied at the same time. As a result, the agents will start behaving selfishly, each trying to maximize their own satisfaction without taking into consideration how their actions might negatively impact others. This context comprises a game where many people, whom we call *agents*, interact with each other in an attempt to satisfy their own desires. These agents each have their own private strategy and their sole aim is to maximize their own "satisfaction" in the final outcome, the so called *utility*.

The field of *Game Theory* tries to formalize, in a principled way, the interactions between the agents, how their actions may lead to different outcomes and the benefits and losses of every agent in those outcomes. In such a game, even if the agents understand game theory, it is not unlikely that if the agents each try to maximize their own payoffs this may lead to a highly inefficient solution, in the sense that in the final outcome, the sum of the players' utilities, the so called *Social Welfare*, may be low. This is the central problem that *Mechanism Design*, the science of rule making, tries to fix: Given that the agents behave selfishly, each trying to maximize their own utility, can we design games, the so called *mechanisms*, where the final solution will be efficient with respect to some objective? For this reason, mechanism design can be thought of as *game theory in reverse*. In this field we mainly focus on designing mechanisms where agents maximize their utilities by acting truthfully, effectively removing the agents' incentives to lie about their preferences. As a result we are able to predict the actions of the players and therefore to reason about the outcome of the mechanism. In simple words, the aim of mechanism design is to design systems for strategic agents that have good performance guarantees.

Informally, a mechanism is described by a set of *feasible outcomes*, $\mathcal{O}$, an *allocation rule*, f, and a payment rule, $\boldsymbol{p} = (p_1, p_2, \dots, p_n)$. Each agent i has a private, i.e. not known to the mechanism, valuation function $v_i(\mathcal{O}) \rightarrow \mathbb{R}_+$, and their goal is to maximize their *utility*, the difference between their value for the final outcome $O \in \mathcal{O}$ minus what they had to pay to the mechanism. In most cases, both the final outcome O and the payment rule are a function of the declared preferences $\boldsymbol{b}$ of all the players, thus agent i's utility is $u_i(\boldsymbol{b}) = v_i(O) - p_i(\boldsymbol{b})$. A mechanism is *truthful* if no agent can increase her utility by misreporting her preferences no matter the actions of the other players in the mechanism. From the definition of the utilities of the agents it becomes apparent that in order to ensure truthfulness, careful consideration of both the allocation and payment rules is required.

The two most standard measures of performance in mechanism design are the total utility of all the agents in the outcome that the mechanism chooses, also known as the *social welfare* of the mechanism, and the total payments that the mechanism collected from all the agents, also known as the *revenue* of the mechanism. The social welfare of the outcome O is formally defined as $\sum_{i=1}^{n} v_i(O)$, while its revenue, when players declare values $\boldsymbol{b}$ to the mechanism is defined as $\sum_{i=1}^{n} p_i(\boldsymbol{b})$. Problems in mechanism design are split into 2 main categories, single-parameter and multi-parameter ones, depending on whether the players declare one or multiple values to the mechanism. In this thesis, we focus on a particular kind of mechanisms: they are called *Auctions* and are mechanisms specifically designed for the exchange of goods and money. Auctions have been extensively studied by both Economists and Computer Scientists.

Single-Parameter Auctions

Suppose you have a single item that you want to auction off and there are n interested potential buyers. From a mechanism design perspective, each of those agents i has a private valuation v_i for the item. If they are the winner of the auction then their utility is their value for the item minus their payment to the auction, i.e. $u_i = v_i - p_i$. If they don't win the auction, then they don't get the item and they obviously pay nothing. In this case, their utility is zero. Let's suppose that the goal of the mechanism designer is to allocate the item to the agent, or bidder in this case, that wants it more, i.e. the one with the highest value for the item. In this setting, this is equivalent to maximizing the social welfare of the mechanism. As we alluded to earlier, whether a mechanism is truthful or not depends on its payment rule. We are going to compare two different payment rules, and illustrate how neither of them achieve truthfulness.

First, let's assume that the mechanism gives the item away for free, i.e. the payment of the winning bidder i is $p_i = 0$, as is everyone else's payment. Let's now observe the utilities of the agents: For every agent i her utility in case she wins the item is now $u_i = v_i$, while her utility in case she doesn't win is zero. It is clear that every bidder has a higher utility for winning the item than not winning it. Therefore, they would misreport their value, declaring to the mechanism as high a value as possible for the item in the hope that they are indeed the winning bidder. Now suppose instead that we charged the winning bidder their declared value, also called her bid, b_i. This mechanism is known as *first-price auction*. Then, the utility of the winning bidder i is $u_i = v_i - b_i$, and the utility of every losing bidder is zero. If a bidder bids her true value and wins the auction, then her utility is zero. The only way for a bidder to have a positive utility in this mechanism is to win the item with a bid lower than her value for it, once again breaking truthfulness. The solution to this problem was given by Vickrey [9], who introduced the aptly named *second-price auction*. In this mechanism, the bidder with the highest bid gets the item, but unlike the first-price auction, now she pays a price equal to the second highest bid. The truthfulness of the second-price auction is presented in subsection 3.2.1.

A question that then arises naturally is the following: If we want to design a truthful mechanism for a different, arbitrary single-parameter environment, or with a different allocation rule, is there a corresponding payment rule such that the resulting mechanism is truthful? The answer to that question was given by Myerson [6], who proved that in order for a mechanism to be truthful its allocation rule must be monotone non-decreasing, meaning that for every bidder, if they increase their bid then their allocation by the mechanism can only increase. At the same time, for a fixed monotone allocation rule, there is a unique payment rule (see subsection 3.2.2) that renders the resulting mechanism truthful. Intuitively, for every quantity that some bidder is allocated, they pay the minimum they could have bid and won that quantity. Myerson's lemma has a very profound consequence: If we are only interested in truthful mechanisms that charge zero to losing bidders, then for any allocation rule, the payment rule is uniquely defined. Therefore, we can think of a mechanism simply in terms of its allocation rule.

Multi-Parameter Auctions

Now assume that we face a more general problem, for example there are multiple identical copies of the item we want to auction off. This setting is known as a *multi-unit auction*. A multi-unit auction consists of a set of m identical units of an item to be allocated to n bidders. Each bidder i has a valuation function $v_i : \mathbb{N} \to \mathbb{R}_+$. It is a natural assumption that valuation functions are non-decreasing, i.e. $v_i(j) \leq v_i(k)$ for all $k \geq j$, and *normalized*, i.e $v_i(0) = 0$. For the objective of social welfare, the goal is to compute an allocation $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ of the units to the bidders that maximizes $\sum_{i=1}^{n} v_i(x_i)$ while for the object of revenue maximization, the goal is to maximize the total payments of the bidders.

In this thesis, we focus on multi-unit auctions with budget-additive or submodular bidders. The valuation function of some bidder i in a multi-unit auction is *budget-additive* if it increases by the same amount for every additional unit they are allocated (up to some maximum amount of units that they are interested in), i.e. their valuation for x_i units is $v \cdot x_i$, where v is their value per unit, and it is *submodular* if for every $j \leq k$ it holds $v_i(j+1) - v_i(j) \geq v_i(k+1) - v_i(k)$. Intuitively, a bidder with additive valuation cares the same for gaining additional units, while a bidder with a submodular one has diminishing returns for multiple units. The class of budget-additive valuations is a proper subset of the class of submodular ones.

Suppose the goal of the mechanism designer is to maximize social welfare. Is there a truthful mechanism that achieves that? Perhaps surprisingly, the answer to that question is positive. For any multi-parameter environment, there is a truthful mechanism that always terminates with an optimal solution. This mechanism was a result of the work by Vickrey [9], Clarke [7] and Groves [8] and is known as the Vickrey-Clarke-Groves (VCG) mechanism. VCG is the unique truthful social welfare-maximizing mechanism and can be applied in very general mechanism design environments. The main idea behind the mechanism is to align the utility of every agent with the social welfare of the final outcome and charge them their *externality*, i.e. the total loss in value they caused to the rest of the bidders (for further details see subsection 3.3.3).

Revenue Maximization

In practice, the one that chooses the rules of the auction is the one who possesses the items that are to be auctioned off, i.e. the auctioneer. The most natural objective for the auctioneer is, undoubtedly, to sell those items for the highest amount of money possible, i.e. *revenue maximization*. For the objective of social welfare, there is a truthful mechanism that in any environment finds the optimal solution. Does something similar hold true for the objective of revenue maximization? The answer is a bit more complicated than it was for social welfare. In case of a single-parameter environment, for all truthful mechanisms, their payment rules are a function of their allocation rules, as defined by Myerson's lemma. Therefore, by taking expectations over the bidders' valuations, for any truthful mechanism in a single-parameter environment the mechanism designer can determine its expected revenue. As it turns out, the expected revenue from every bidder is equal to a function of the distribution of their valuation. This function is called their *virtual valuation function*. Then, the problem of revenue maximization reduces to simply choosing the allocation rule that maximizes the expected revenue, i.e. the sum of the virtual values of all the bidders. The resulting mechanism is named Myerson, and under some mild assumptions about the bidders' valuation distributions, is truthful and maximizes the expected revenue. For general multi-parameter environments, the problem is even more complicated. Now Myerson's lemma no longer holds, thus reasoning about the payments of any mechanism is a much more involved process. Even for two bidders and two items available for sale, the revenue-maximizing mechanism is not known.

Automated Mechanism Design

Maximizing revenue in multi-parameter environments is particularly challenging. Even for relatively simple settings, revenue-maximizing mechanisms are in most cases not known. Furthermore, for the settings where the revenue-maximizing mechanism is known, in order for it to work it requires knowledge of the distribution over the bidders' valuations. In an attempt to circumvent these two fundamental issues, the idea of using samples to guide the design of mechanisms was born. Sample-based mechanism design was first introduced in the context of *automated mechanism design*. In automated mechanism design, the goal is to design algorithms that take as input information about the set of bidders and return a mechanism that maximizes some objective function, such as revenue [32, 33, 34]. In the early days of automated mechanism design, the input information about the bidders was an explicit description of the distribution over their valuations. The support of this distribution is often doubly exponential, rendering obtaining and storing it prohibitive. In response, sample-based mechanism design was introduced where the input is now a set of samples from this distribution [35, 36, 37].

Another important contribution of these papers was that they introduced the idea of searching for the best mechanism in a parameterized space where any parameter vector yields an incentive-compatible mechanism, meaning that every bidder has an incentive to participate in the mechanism and bid truthfully. This was in contrast to the traditional, up to that point, approach of viewing mechanism design as an unrestricted optimization problem and modeling those constraints afterwards.

This parameterized approach to mechanism design became very popular in the research community. Numerous follow-up papers adopted this approach of parameterized, sample-based revenue maximization, while also trying to prove generalization guarantees for their results [23, 38, 25, 39]. The aim in these papers is not simply to learn some parameter vector such that the corresponding mechanism has high revenue in the samples, but to provide *generalization guarantees* that this mechanism will also perform well in the real world, provided that the samples were independent draws from the true distribution of the bidders' valuations. This statistical guarantee usually takes the form: "With probability at least $1 - \delta$, the expected revenue of the mechanism we determined on the samples will be at most ϵ lower in the real world than it was in the samples, provided we used at least a number N of samples". The mechanism designer then has the choice to configure both the 'chance of failure' δ and the expected 'generalization error' ϵ according to her needs, and the required number of samples N changes accordingly. Because these results give the required number of samples for some generalization guarantee to hold, they are called *sample complexity* results in the literature.

Two of the most impactful papers in this direction were [23] by Morgenstern and Roughgarden and [12] by Balcan et al. In the first paper, Morgenstern and Roughgarden introduced the class of t-level auctions. Intuitively, every auction in that class uses, for every bidder, a step function with t steps as a proxy for their virtual valuation function, and chooses a final allocation that is optimal with respect to these steps. The authors then proved that, for what they characterized as 'simple' environments, a close approximation of those step functions to the virtual valuation functions of the bidders implies an equally close approximation of the expected revenue of the corresponding auction to that of the optimal one. Then, for these settings, the problem of maximizing revenue reduces to learning a t-step function for every bidder that approximates closely enough their virtual valuation function. They then upper bounded the number of required samples, i.e. the sample complexity to do so. This means that for those settings, they upper bounded the sample complexity of learning revenue-optimal auctions. The second paper was much less applied. In that paper, the authors didn't propose any new mechanism classes but they instead introduced a novel way of reasoning about the expressive capability, and therefore the intrinsic complexity, of any mechanism class. As we discussed earlier, in the parameterized approach to automated mechanism design, any mechanism from a given mechanism class is uniquely defined by some parameter vector $\boldsymbol{p}$. Therefore, if we fix some arbitrary valuation profile $\boldsymbol{v}$, then the revenue of any mechanism in that class is solely a function of its parameter vector $\boldsymbol{p}$. Balcan et al. proved that if, for any fixed $\boldsymbol{v}$, the parameter space of the class can be broken into regions where the revenue on that valuation profile is a linear function of $\boldsymbol{p}$, then the pseudo-dimension of that mechanism class can be directly upper bounded in relation to the number of such regions. Pseudo-dimension is a measure of the expressive capability of any class. Using that measure, one can then easily bound the sample complexity of determining an optimal mechanism from that class. Effectively, with this paper the authors provided a framework for reasoning about the sample complexity of any mechanism class, provided that the mechanism designer has a very good understanding of the class in question.

Deferred-Acceptance Auctions

Now suppose that you want to design a mechanism for some environment with very complex restrictions on the set of feasible solutions. Even for the seemingly easy objective of social welfare maximization, calculating the optimal solution can take doubly exponential time. A logical next step then is to field a mechanism that uses some greedy algorithm to choose a final solution. Greedy algorithms have been used extensively in mechanism design. They may not find the optimal solution, but for many settings, they can achieve a constant approximation of it. Milgrom and Segal [14] took this idea of using greedy algorithms one step further. Traditional greedy approaches to mechanism design were forward-greedy: The allocation algorithm one by one accepted the bidders that seemed the most 'promising', according to its greedy heuristic and subject to feasibility constraints, until a maximal feasible solution was reached. Milgrom and Segal instead proposed the framework of *deferred-acceptance auctions* (DA), a family of mechanisms that can be thought of as running an adaptive, backward greedy algorithm for deciding the set of accepted bidders: Now at every round the 'least promising' bidder is excluded by the mechanism, until the set of bidders that remain constitute a feasible solution.

The fact that DA auctions use backward greedy algorithms instead of forward ones gives affords them many appealing properties. For one, DA auctions are truthful, even in the sense of *obvious strategyproofness* formalized by Li [16]. This means that even non-expert bidders can understand that truthful bidding is a dominant strategy, and in turn, will be incentivized to participate in the mechanism. This is crucial in settings where the participation of more agents in the mechanism enables it to find better solutions. Furthermore, every DA auction satisfies *group-strategyproofness*. This means that no coalition of bidders can collectively submit false bids in such a way that makes every bidder of that coalition strictly better off. Effectively, no bidder has an incentive to collude against the auctioneer. For those reasons DA auctions have been fielded in many critical applications, namely for the reverse auction part of spectrum reallocation [17]. DA auctions, as they were introduced by Milgrom and Segal, were restricted to binary environments, where every bidder is either a 'winner' or a 'loser'. Gkatzelis, Markakis and Roughgarden [18] extended the framework of DA auctions to non-binary settings, where each bidder receives some level of service, subject to feasibility constraints.

The performance of DA auctions for the objective of social welfare has been studied extensively. Unfortunately, in some settings DA auctions do not achieve a good approximation for that objective. A high-level explanation would be that in those settings, backward greedy algorithms are inferior to forward greedy ones, and the former may not even lead to maximal solutions. Motivated by this concern, Dütting et al. [20] explored the performance and limitations of deferred-acceptance auctions for the objective of social welfare maximization from an approximation algorithms viewpoint. On the positive side, for combinatorial auctions with single-minded bidders DA auctions can nearly match the performance of arbitrary truthful and computationally efficient mechanisms. On the negative side, even for 2 bidders and 2 copies of the same item available for sale, with arbitrary valuations no DA auction can guarantee an approximation factor to social welfare that is better than $\frac{1}{\sqrt{2}}$ [18].

Our Results: Intuition and Contribution

The performance of deferred-acceptance auctions for the objective of social welfare maximization has been well documented. On the contrary, the power and limitations of DA auctions for the objective of revenue maximization are much less understood. In this thesis, our aim is to discover environments in which, under some natural distribution assumptions, deferred-acceptance auctions perform well for the objective of revenue maximization. Specifically, we determine environments in which, given some samples of the bidders distributions, it is possible to learn DA auctions with high expected revenue. With this goal in mind, we explore both single- and multi- parameter environments.

The first single-parameter environment we focus on are multi-unit auctions with budget-additive bidders. In this environment, there are identical copies of the same good available for sale and the bidders have a constant value per unit of the good that they are allocated. This is true up to the point where they are no longer interested in acquiring more units. They each declare a single number to the mechanism, their valuation per unit of the good. Intuitively, in this environment, VCG is implementable as a DA auction [18]. Therefore, if every bidder's virtual valuation function was known then the revenue-optimal mechanism, as the virtual surplus-maximizing mechanism, would also be implementable as a DA auction. Thus, the problem of learning a revenue-optimal DA auction in this environment reduces to learning the virtual valuation function of every bidder. To do this, it suffices to observe that in this environment, t-level auctions are implementable as DA auctions. After some slight modifications to their pseudo-dimension argument (because it doesn't carry over to multi-unit environments immediately), we have successfully upper bounded the sample complexity of learning revenue optimal t-level multi-unit auctions for budget-additive bidders, which are also DA auctions. Encouraged by this result, we then moved to single-parameter environments with arbitrary polymatroid constraints. Pictorially, now there are no indivisible units to be allocated to the bidders, but instead the space of feasible allocations constitutes some polyhedron in the n-dimensional space. This environment is highly expressive, and can capture very intricate restrictions on the set of feasible restrictions. A first idea, since this is still a single-parameter, albeit complex, environment, would be to once again use t-level auctions. However, this approach is flawed: The final allocation of a t-level auction is optimal with respect to its levels, not the actual value of the corresponding virtual valuation function at those points. The problem is that those levels are not proportional to the actual value of the virtual valuation function at those points. Therefore, maximizing with respect to the levels, for complex environments such as those with polymatroid constraints, does not necessarily mean that the solution is also optimal with respect to the virtual valuation functions of the bidders. To circumvent this, we proposed a new mechanism class, which we named *extended t-level auctions*. The difference is that now every level is supplied with one additional parameter, its *level score*. The auction does not choose a solution that is maximal with respect to the levels, but with respect to those level scores. Intuitively, if the level scores at every level are proportional to the virtual valuation function of the corresponding bidder at that point, then the level score-optimal allocation will also be approximately optimal with respect to the virtual valuation functions of the bidders. This mechanism class has significantly higher expressive capabilities than the original t-level auctions, yet we proved that the same pseudo-dimension and sample complexity bounds as for the original t-level auctions still hold (theorems 5.7 and 5.8).

Then we shifted our attention to multi-parameter environments. Specifically, we studied multi-unit auctions with submodular bidders. Again, there are m identical copies of the same item available for sale, but now the bidders experience diminishing returns for additional copies of the same item. They each declare to the mechanism m numbers, their value for the first unit, then their additional value for a second, a third and so on. This environment introduces two new challenges: Now the revenue-optimal mechanism is not known, which means that we will have to compare the expected revenue of any mechanism we propose directly against the expected optimal social welfare, which is an immediate upper bound on the expected revenue of any individually rational mechanism. Furthermore, in this environment VCG is not implementable as a DA auction [18, Section 7]. This means that even the expected social welfare of the optimal mechanism for that objective is not very close to the expected optimal social welfare, let alone the expected revenue. Our approach to this environment was quite novel: Intuitively, the issue with 'conventional' DA auctions for this environment is that for subsequent units, DA auctions need to charge a higher price in order to ensure truthfulness. When bidders are submodular however, their value per unit decreases. This means that either the prices of the initial units will be too low, and a lot of potential revenue will be 'left on the table' for the bidders, or the prices will be too high initially and a lot of units will remain unsold. In order to get around this problem, we propose the *unit bundling mechanism class*: Each bidder is offered a specific bundle of units, at a specific total price, and they can either accept or decline their offer. The parameters of any mechanism from that class are the bundle size of each bidder and the corresponding asking price for that bundle. In order to compare the expected revenue of the optimal mechanism from this class against the optimal social welfare, we use the expected social welfare of the optimal fixed allocation as an interface: We compare separately the expected revenue of the optimal mechanism from the unit bundling class against the expected social welfare of the optimal fixed allocation and then we compare the expected social welfare of that allocation against the expected social welfare of VCG. After this the only thing left to show is that it is possible to learn those optimal bundle sizes and reserve prices. To do this, we proved that the parameter space can be broken into regions where the revenue is linear and used Balcan's sample complexity framework to bound the pseudo-dimension of the class. Putting these together, we upper bounded the sample complexity of learning a DA auction that, provided the bidders' distributions are Sub-Gaussian, achieves expected revenue close to the expected Social Welfare of VCG, minus some small additive terms.

Finally, we studied how these results change in the case that the number of samples is severely restricted. In the case of a multi-unit auction with budget-additive bidders, we were heavily inspired by [29]. In that paper, Hartline and Roughgarden introduced the notion of *commensurate mechanisms* for binary environments. They proved that if a mechanism is commensurate with another, then the first achieves at least half the expected revenue of the second. In this thesis we extended the notion of commensurate mechanisms to arbitrary single-parameter environments in the most natural way, and similarly proved that the same revenue property still holds. Using this notion of commensurate mechanisms, we introduced a new DA auction that, using a single sample, achieves on expectation at least one fourth of the expected revenue of Myerson. For multi-parameter environments, our proposed mechanism can be described quite simply: Use one sample to determine an approximation of the optimal bundle sizes for every bidder and a second one to determine the asking prices. The proof that the resulting bundle sizes using a single sample will be close to optimal with respect to social welfare was done from scratch. The proof that those reserve prices are close to the optimal ones was an application of one of the main results of [31]. The resulting mechanism achieves on expectation, using only two samples, an almost 0.589-approximation of the expected revenue of the optimal mechanism from the unit bundling class.

Organization of the Thesis

In chapter 3 we will make a brief introduction to *Mechanism Design*, going over fundamental definitions and theorems. We will overview the notions of *dominant strategy*, *truthfulness* and the *Revelation Principle*. For single-parameter environments, we will explain the *Second-Price Auction*, *Myerson's Lemma* and the *Myerson Mechanism*. For multi-parameter environments we will overview the basic valuation classes and their relationship as well as the VCG mechanism.

In chapter 4 we will introduce the reader to some of the most recent trends in mechanism design. In the first half of this chapter we will make a brief introduction to PAC learning and the notion of uniform convergence and how these can be applied to automated mechanism design. Then we will overview the notions of generalization and representation error bounds, and explain how these two can be used together to establish what are called sample complexity bounds. Finally, we will present the main results from Balcan's sample complexity framework [12]. In the second half of this chapter we will present deferred-acceptance auctions in more detail. We will formalize their definition and their remarkable incentive properties and overview their power and limitations.

All our contributions are included in chapters 5 to 7. In chapter 5 we study single-parameter environments. First, we show how for multi-unit auctions with budget-additive bidders t-level auctions are implementable as DA auctions, which means that we immediately get an upper bound on the sample complexity of learning revenue-maximizing DA auctions for this environment. Then, we introduce the class of extended t-level auctions and upper bound the sample complexity of learning revenue optimal auctions for environments with polymatroid constraints using this mechanism class. In chapter 6 we introduce the *a priori optimal mechanism* and the *unit bundling* mechanism class. We show that this mechanism has expected revenue close to that of VCG. Then, using Balcan's framework to bound the pseudo-dimension of the unit bundling class and the a priori optimal mechanism as an interface between that class and VCG, we upper bound the sample complexity of learning auctions from that class with high expected revenue. Finally, in chapter 7 we explore how the results of both of the previous chapters change in case where the number of samples is severely restricted.

Chapter 3

Basics of Mechanism Design

First we will present some fundamental definitions in the area of mechanism design as well as the most basic mechanisms for single and multi-parameter environments.

3.1 Preliminaries

The basic setup: Assume that there are n agents participating in the mechanism, and let $\mathcal{O}$ be the set of feasible outcomes. Each agent i has a private valuation function $v_i : \mathcal{O} \to \mathbb{R}_+$, representing her value for each possible outcome, and reports her bid $b_i : \mathcal{O} \to \mathbb{R}_+$ to the mechanism. After collecting the bids, the mechanism uses a function $f : \boldsymbol{b} \to \mathcal{O}$, called the *allocation rule*, which maps the *bid profile*, $\boldsymbol{b} = (b_1, b_2, \dots, b_n)$, to an outcome, and a function $p : \boldsymbol{b} \to \mathbb{R}^n_+$, called the *payment rule*, which determines the payments of all bidders. Any deterministic mechanism is uniquely defined by its pair of functions (f, p).

Notation: By $\boldsymbol{x}_{-i}$ we express the vector $\boldsymbol{x}$ with its $i-$th coordinate removed, i.e.
$\boldsymbol{b}_{-i} = (b_1, b_2, \dots, b_{i-1}, b_{i+1}, \dots, b_n)$

Definition 3.1. ***(Quasi-Linear Utility, [3]).*** *In a mechanism (f, p) we say that the utility functions of all players are quasi-linear if:*

$$u_i = v_i(f(\boldsymbol{b})) - p_i(\boldsymbol{b}) \ \ \forall i \in [n] \tag{3.1}$$

where $\boldsymbol{b}$ is the bid profile of the agents and $[n] = \{1, 2, \dots, n\}$.

Note: In all of our work, we assume that all agents have quasi-linear utilities. This is standard in mechanism design.

Definition 3.2. ***(Dominant Strategy).*** *A bidding strategy b_i is dominant if it maximizes agent i's utility, regardless of the strategies $\boldsymbol{b}_{-i}$ of the other agents. Formally:*

$$v_i(f(b_i, \boldsymbol{b}_{-i})) - p_i(b_i, \boldsymbol{b}_{-i}) \geq v_i(f(b'_i, \boldsymbol{b}_{-i})) - p_i(b'_i, \boldsymbol{b}_{-i}) \tag{3.2}$$

for every other strategy b_{-i} of agent i, and every strategy profile b_{-i} of the other agents.

A mechanism is called *truthful* if for every agent, truthtelling is a dominant strategy. Formally, we have:

Definition 3.3. ***(Truthfulness, [2]).*** *Let (f, p) be a mechanism. Then, mechanism (f, p) is truthful if for all $i \in [n]$ and for every v_i' it holds that:*

$$v_i(f(v_i, \boldsymbol{v}_{-i})) - p_i(v_i, \boldsymbol{v}_{-i}) \geq v_i(f(v_i', \boldsymbol{v}_{-i})) - p_i(v_i', \boldsymbol{v}_{-i}) \tag{3.3}$$

When an agent enters a mechanism, it is important for her to know that her participation in the mechanism cannot result in her having negative utility. Mechanisms satisfying this property are called *individually rational*. Formally:

Definition 3.4. ***(Individually Rational, [2]).*** *A mechanism (f, p) is individually rational if for all $i \in [n]$ and for all valuation profiles $\boldsymbol{v} = (v_1, v_2, \dots, v_n)$, it holds:*

$$v_i(f(\boldsymbol{v})) - p_i(\boldsymbol{v}) \geq 0 \tag{3.4}$$

Definition 3.5. ***(Dominant Strategy Incentive Compatible).*** *A mechanism is Dominant Strategy Incentive Compatible (DSIC) if it is truthful and individually rational.*

Assume that some mechanism (f, p) is non-DSIC, in the sense that every agent has a dominant strategy, but that strategy might not be truthtelling. The question that naturally arises is whether (f, p) can be simulated by another mechanism, (f', p'), such that (f', p') is DSIC. The answer to that question is positive, and this result is called the *Revelation Principle*. Formally, we have:

Theorem 3.1 (Revelation Principle)**.** *For every mechanism (f, p) in which every agent has a dominant strategy, there is an equivalent direct-revelation DSIC mechanism (f', p').*

Proof. From our assumptions, every agent has a private valuation v_i and a dominant strategy $s_i(v_i)$. This means that each agent would declare her valuation as $s_i(v_i)$ to the mechanism (f, p). We can now construct an equivalent mechanism (f', p') that accepts each agent's bid b_i, applies the function s_i on the declared bid b_i for every agent $i \in [n]$ and then outputs the same allocation and payments as (f, p) on those transformed bids. Formally, we can define $f'(\boldsymbol{b}) = f(s(\boldsymbol{b}))$ and $p'(\boldsymbol{b}) = p(s(\boldsymbol{b}))$, where $s(\boldsymbol{b}) = (s_1(b_1), s_2(b_2), \dots, s_n(b_n))$.
As a result, an agent i who has private valuation v_i and dominant strategy $s_i(v_i)$ would only reduce her utility by reporting a bid other than $s_i(v_i)$ pm (f, p), and therefore a bid other than b_i on (f', p'). Hence, mechanism (f', p') is DSIC. □

Up to this point, we have not mentioned anything regarding the measure of efficiency of mechanisms. Naturally, efficiency varies according to the mechanism designer's wishes. For example, if a government is to decide about the construction or not of a public project, then the appropriate efficiency measure is the total welfare of the community. If a seller was to auction her car, she probably would want to earn as much money from the sale as possible. In this case, the appropriate efficiency measure would be the revenue of the mechanism. Formally:

Definition 3.6. ***(Social Welfare).*** *Let (f, p) a mechanism and $o \in \mathcal{O}$ its output on some valuation profile. Then, the Social Welfare of that outcome is defined as:*

$$SW = \sum_{i=1}^{n} v_i(o) \tag{3.5}$$

The revenue of the mechanism is defined naturally as the sum of the payments made to the mechanism by all participating agents. Formally:

Definition 3.7. ***(Revenue).*** *Let (f, p) be a mechanism and $\boldsymbol{b} \in \mathbb{R}_+^n$ the bid profile of the agents. Then, the revenue on this bid profile is defined as:*

$$REV = \sum_{i=1}^{n} p_i(\boldsymbol{b}) \tag{3.6}$$

Finally, we are going to define the concepts of approximation and randomized algorithms, which are widely used in Mechanism Design:

Definition 3.8. ***(Approximation in Mechanism Design, [3]).*** *We say that a mechanism ρ-approximates the optimal solution if: $ALG \geq \rho \cdot OPT$, where $\rho \leq 1$, ALG is the efficiency of the mechanism and OPT is the efficiency of the optimal solution, for the given efficiency measure.*

Definition 3.9. ***(Randomization in Mechanism Design,[3]).*** *We say that a randomized mechanism ρ-approximates the optimal solution if: $\mathbb{E}[ALG] \geq \rho \cdot OPT$, where $\rho \leq 1$, ALG is the efficiency of the mechanism and OPT is the efficiency of the optimal solution, for the given efficiency measure.*

3.2 Single-Parameter Environments

3.2.1 Single-item Auctions

Suppose that we want to design an auction for an indivisible item, so as to maximize the Social Welfare, i.e. allocate the item to the agent that wants it the most. As hinted earlier, we cannot give the item for free, nor charge the winning agent her bid, as in both these scenarios agents have incentives to misreport their true valuations, and this can lead to inefficient allocations. Thus, we have to find the appropriate payment rule, so that no agent can increase her utility by lying about her valuation. The solution to this problem is known in the literature as the *Second-Price* or *Vickrey Auction*: In this auction, every bidder submits her bid. Once all bids are collected, the highest bidder gets the item, and pays for it a price equal the second highest bid.

Theorem 3.2. *The Second-Price Auction is DSIC.*

Proof. Let n be the number of agents that participate in the auction. Fix an arbitrary agent i, and let v_i be her valuation for the item for sale. It suffices to prove that it is a dominant strategy for bidder i to bid her true valuation, i.e. $b_i = v_i$. Let $\boldsymbol{b}_{-i}$ be the bid profile of the other agents, and let B be the highest bid among them, i.e. $B = \max_{j \neq i} b_j$. There are 2 distinct cases:

- $v_i < B$: If $b_i < v_i$ then the outcome of the auction remains the same as if agent i had bid v_i, and she does not get the item. Her utility then remains zero, as it would have been if she had instead bid her true valuation. If $b_i > v_i$ there are 2 possible outcomes: If it still holds that $b_i < B$ bidder i again doesn't win the item and her utility remains 0. If on the other hand it holds that $b_i > B$ then bidder i will win the item, and pay a price equal to the second highest bid, B. Now her utility is $v_i - B < 0$. Thus, if $v_i < B$, agent i cannot increase her utility by misreporting her valuation.
- $v_i \geq B$: If $b_i > v_i$ then the outcome of the auction remains exactly the same as if bidder i had bid her true valuation: Bidder i wins the item, and she pays a price equal to the

second highest bid, B. If on the other hand $b_i < v_i$ there are 2 possible outcomes: If it still holds that $b_i > B$ then bidder i still wins the unit, and pays a price of B, so her utility is unchanged. If however $b_i < B$, then bidder i no longer wins the item, and she pays nothing.

In any case, it is clear that agent i cannot increase her utility by misreporting her valuation, regardless of the bids of the other agents. Thus, truthtelling is a dominant strategy and the mechanism is DSIC. □

3.2.2 Multi-Unit Auctions and Myerson's Lemma

Now suppose that we have multiple, or even infinite copies of the same item available for sale. These auctions are called *multi-unit auctions*. This time agents report their "valuations per unit" of the good. A modification of the second-price auction could also work in this environment. A question that naturally arises in practice is the following: If we want to implement a different allocation rule $f(\cdot)$, is there a corresponding payment rule $p(\cdot)$ that would make the resulting mechanism (f, p) DSIC? The answer to this question was given by *Myerson*. Formally:

Theorem 3.3. *(Myerson's Lemma, [6]). A mechanism (f, p) for a single-parameter environment is DSIC if and only if, for every bidder i with bid b_i and bid profile $\boldsymbol{b}_{-i}$ by the other players, it holds:*

- *$f_i(b_i, \boldsymbol{b}_{-i})$ is non-decreasing in its first argument.*
- *There is a unique payment rule making the mechanism (f, p) DSIC. That rule is given by the formula:*

$$p_i(b_i, \boldsymbol{b}_{-i}) = \int_0^{b_i} z \frac{d}{dz} f_i(z, \boldsymbol{b}_{-i}) dz \tag{3.7}$$

3.3 Multi-Parameter Environments

In the previous section, we restricted our attention to cases where bidders only reported a single number to the mechanism. Informally, this is the definition of *single-parameter environments*. But what if we are facing a more complicated problem, e.g. one where there are multiple different items to be auctioned, and bidders have different valuations for different sets of items? What would an appropriate payment rule in this case look like, in order to ensure truthfulness? A reasonable assumption would be that running a separate *Second-Price Auction* for each item would result in a DSIC mechanism, since as we discussed earlier, the *Second-Price Auction* is DSIC. This is true if bidders have *additive valuations*, meaning that for every bidder, her valuation for a set of items is simply the sum of that bidder's valuations for the items contained in that bundle. However that is not generally the case: For complex bidder valuations, running separate *Second-Price Auctions* for each item is not a DSIC mechanism. It is apparent that when designing a mechanism in such settings, careful consideration of the bidders' valuations is required.

3.3.1 Combinatorial Auctions

We are now ready to define the notion of a *Combinatorial Auction*. Informally, these are a special case of a *multi-parameter environment* defined earlier, where the valuation function of every bidder is defined over the set of all 2^m subsets of the m different items. In this case, each different outcome is uniquely defined by specifying for every item, which bidder, if any, gets it. Thus, the sets of possible outcomes is exactly $(n+1)^m$, and different from the set of values that every bidder must report to the mechanism. The reason for this differentiation behind the set of declared values and possible outcomes is that every agent is indifferent to who gets an item, if it is not her.

Definition 3.10. ***(Combinatorial Auction).*** *A Combinatorial Auction consists of a set U of m distinct items to be allocated to n bidders. Each bidder i has a valuation function $v_i : 2^U \to \mathbb{R}_+$. Valuations functions v are assumed to be normalized, i.e. $v(\emptyset) = 0$, and non-decreasing, i.e. $v(S) \leq v(T)$ for all $S \subseteq T$. This last property is called free disposal.*

3.3.2 Valuation Classes

In this section we are going to give some basic definitions about valuation classes. Valuation classes are strongly correlated with the difficulty of the mechanism design problem. Let v be a valuation function, and U the set of m, possibly distinct, items.

Definition 3.11. ***(Additive Function).*** *A set function $v : 2^U \to \mathbb{R}_+$ is additive if for every $S \subseteq U$:*

$$v(S) = \sum_{j \in S} v(\{j\}) \tag{3.8}$$

This is the least general class of valuation functions and implies that there are no dependencies between the items of any possible set. As discussed earlier, if the mechanism design goal is to maximize social welfare, in this setting the problem can be solved optimally through parallel second-price auctions.

Definition 3.12. ***(Gross Substitutes Valuation).*** *An agent is said to have a gross substitutes valuation if, whenever the prices of some items increase and the prices of other items remain constant, the agent's demand for the items whose price remain constant weakly increases.*

The above definition is rather informal. However, this thesis does not involve gross substitutes valuations, therefore the formal definition is deferred since it requires some more technical background.

Definition 3.13. ***(Submodular).*** *A set function $v : 2^U \to \mathbb{R}_+$ is submodular if for every $S \subseteq T \subseteq U$ and item $j \notin S$ it holds:*

$$v(S \cup \{j\}) - v(S) \leq v(T \cup \{j\}) - v(T) \tag{3.9}$$

Submodularity can be seen as the discrete analog of concavity. A more intuitive interpretation is that agents have "diminishing returns" for being allocated more items. Submodular valuations arise naturally in many economic settings.

Definition 3.14. ***(Fractionally Sabadditive (XOS)).*** *A set function $v : 2^U \to \mathbb{R}_+$ is fractionally subadditive if there exist additive functions $w_k : 2^U \to \mathbb{R}_+$ such that for every $S \subseteq U$:*

$$v(S) = \max_k w_k(S) \tag{3.10}$$

Definition 3.15. ***(Subadditive).*** *A set function $v : 2^U \to \mathbb{R}_+$ is subadditive if for each subset S and T of U:*

$$v(S) + v(T) \geq v(S \cup T) \tag{3.11}$$

If all agents participating in a market have subadditive valuations, then the combination of any two bundles of items cannot increase their value. This is also called a *complement-free market.*

The main reason we mentioned these valuation classes is the natural relation between them:

$$\text{Additive} \subseteq \text{Gross Substitutes} \subseteq \text{Submodular} \subseteq \text{XOS} \subseteq \text{Subadditive} \tag{3.12}$$

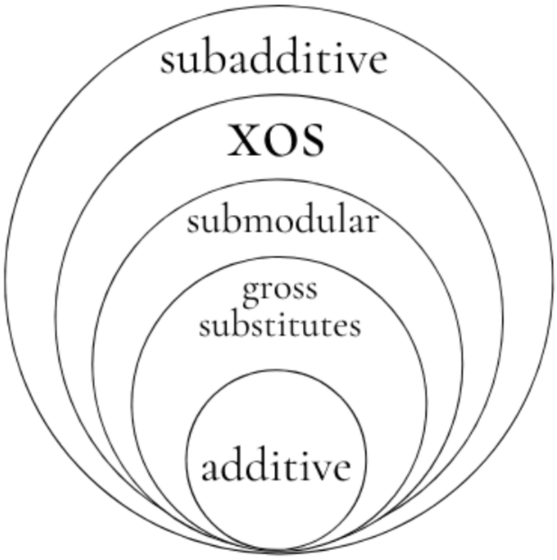

Figure 3.1: Relation between valuation function classes

3.3.3 Vickrey-Clarkes-Groves Mechanism

Based on the diversity of possible valuation classes of the agents, one may assume that in complex settings with general valuation functions maximizing the Social Welfare might be impossible. Perhaps surprisingly, there exists a DSIC mechanism such that in any setting and for any valuation profile, it maximizes the Social Welfare. To illustrate this mechanism, we'll first define a multi-parameter environment in its most abstract form:

- There are n agents.
- There is a set $\mathcal{O}$ of possible outcomes.
- Every agent has a private valuation function $v_i(o), \forall o \in \mathcal{O}$.

In any such environment, there is a DSIC mechanism that maximizes Social Welfare when the agents utilities are quasi-linear. This mechanism is called the Vickrey-Clarkes-Groves mechanism (VCG). Formally:

Theorem 3.4. ***(VCG mechanism, [7, 8, 9]).*** *In every general mechanism design environment, there is a DSIC Social Welfare-maximizing mechanism (f, p) with allocation and payment rules defined as follows:*

- $f(\boldsymbol{b}) = o^* = \arg\max_{o \in \mathcal{O}} \sum_{i=1}^{n} b_i(o)$
- $p_i(\boldsymbol{b}) = max_{o \in \mathcal{O}} \sum_{j \neq i} b_j(o) - \sum_{j \neq i} b_j(o^*)$

Intuitively, each agent i is asked to pay the difference that her participation in the mechanism caused to the Social Welfare of all the other agents except her. This is known as the *externality* of agent i.

Proof. Fix an arbitrary agent i, and let v_i be her valuation. Our first goal is to prove that it is a dominant strategy for agent i to bid her true valuation function, i.e. $b_i = v_i$. Let $\boldsymbol{b}_{-i}$ be the bid profile of all the other agents except i. From the definition of the allocation and payment rules, it follows immediately that the utility of bidder i is:

$$u_i = v_i(o^*) - p_i(o^*) = \left[v_i(o^*) + \sum_{j \neq i} b_j(o^*) \right] - \max_{o \in \mathcal{O}} \sum_{j \neq i} b_j(o) \tag{3.13}$$

Agent i cannot influence the final term of the right hand side of equation (3.13), because that term only depends on the bids of the other agents. Therefore, in order to increase her utility, agent i should maximize the first of the right hand side of equation (3.13). However, by the definition of the allocation rule, the term $\left[v_i(o^*) + \sum_{j \neq i} b_j(o^*) \right]$ is maximized when agent i reveals her true valuation profile. Misreporting can only lead to outcomes with lower welfare, and lower utility for agent i. Thus, it is a dominant strategy for every agent i to bid her valuation function. Finally, it is easy to verify that truthtelling guarantees non-negative utilities for all agents, since the maximization space for the first term of the RHS of equation (3.13) is strictly bigger than that of the negative term. Thus, the VCG mechanism is DSIC. $\square$

Despite its nice theoretical properties, VCG is rarely used in practice. There are many valid reasons behind VCG's obscurity. Suppose we need to create a combinatorial auction for some setting with n agents and m items available for sale. For m items, there are 2^m different possible item bundles. VCG can indeed find the socially optimal solution, but in order to do so, each agent must declare her value for each of those bundles. Therefore, each agent needs to calculate, and communicate to the mechanism, her value for 2^m bundles. It follows immediately that the amount of values that each agent needs to communicate to the mechanism grows exponentially with the number of items available. And this is a *communication bound*: It holds even if $P = NP$.

But VCG does not suffer only from the communication problem that we just described: It can only be implemented in polynomial time for very specific classes of valuations functions. In general environments, the time required for implementing VCG grows exponentially with the number of goods available in the auction.

Finally, even though we proved that VCG is DSIC, meaning that for every agent it is a dominant strategy to bid her true valuation, VCG is vulnerable to *collusion* among agents: This means that if a group of agents cooperated, they could possibly all misreport their valuations in such a way that strictly increased the utility of every agent in the colluding group. The interested reader can learn more in [1].

Example 3.1. *Let's take as an example a case where there are* 3 *agents, and the municipality is trying to decide whether to build a school or a public pool in some plot of land. Agents 1 and 2 have are parents and don't like swimming very much, so it is* $v_1(school) = v_2(school) = 5$ *and* $v_1(pool) = v_2(pool) = 0$*. Agent* 3 *on the other hand has no children, but she is an avid swimmer:* $v_3(pool) = 9$ *and* $v_3(school) = 0$*. If all agents bid their true valuations, then VCG would decide that a school should be built, and the payments would be* $p_1 = p_2 = 4$ *and* $p_3 = 0$*. If however, agents 1 and 2 colluded and reported* $v_1'(school) = v_1'(school) = 10$ *then again, the VCG mechanism would decide that a school should be built, but now the payments would instead be* $p_1' = p_2' = p_3 = 0$*. Perhaps surprisingly, by bidding more, agents* 1 *and* 2 *and up paying less.*

This example illustrates one more potential issue with VCG: Even for relatively simple environments, VCG may exhibit bad *revenue performance*. In some cases, like a public project or any other auction organized by the public sector, the revenue of the mechanism may not be a major concern. In most practical applications however, the *expected revenue* is one of the primary factors driving mechanism design.

3.4 Revenue Maximization and Bayesian Analysis

In the previous section we saw that even though VCG maximizes social welfare, we cannot provide strong guarantees about its revenue. One may wonder why we started our analysis with the goal of welfare-maximization, and not the perhaps more natural objective of revenue-maximization. This is a common convention in most Algorithmic Game Theory courses, and the main reason behind this choice is pedagogical: social welfare is special. In every single or multi-parameter environment, there is a DSIC mechanism for maximizing social welfare *ex post*, intuitively meaning as well as if the auction designer knew the private valuations of all agents in advance. This is a very strong performance guarantee, and it cannot generally be achieved for most other possible objective functions. This is also the case for revenue maximization.
The following example is illuminating. Suppose there is only one agent, and one item available for sale. In this case, the space of DSIC, direct-revelation mechanisms is very simple: The auctioneer posts a fixed price, also called the *reserve price* for the item. If the bidder's valuation for the item is higher than that price, then she buys the item at that price, else the item remains unsold. This family of mechanisms is called *posted-price mechanisms*.
How can we reason about their revenue performance? If the posted price is higher than the value of the agent for the item, then she won't buy it and the revenue of the mechanism is trivially zero. If on the other hand the posted price is lower than the value of the agent for the item, then she will buy it, and the revenue of the mechanism is exactly equal to the posted price. Therefore, as long as the agent buys the item, the higher the posted price, the higher the revenue. It becomes apparent that unlike the objective of welfare maximization, for the objective of revenue maximization, different auctions do better on different inputs.

3.4.1 Bayesian Analysis

For this reason, comparing different auctions for revenue maximization requires a model to reason about trade-offs across different inputs. The most prevalent and well-studied model for doing this is *average-case* or *Bayesian analysis*. The crux of the matter is that every bidder i's valuation is a sample from some distribution F_i. The auctioneer doesn't know the private valuations of the bidders, but knows their distributions. In practice, these distributions could

have been derived from data from past auctions. The aim of the auctioneer now is to maximize the *expected revenue* of the auction.
Let's focus on single-parameter environments. If we restrict our attention to DSIC, direct-revelation mechanisms then from Myerson's lemma we know the unique payment rule p for any monotone allocation rule f. Therefore, the problem of revenue maximization for single-parameter environments reduces to finding the allocation rule that maximizes the expected payments according to Myerson's Lemma, and ensuring that that allocation rule is indeed monotone.
If we assume truthful bids (because we only consider DSIC mechanisms), express the revenue of the mechanism as the sum of expected payments according to Myerson's lemma and take linearity of expectations over the valuations of the players, after some manipulation of the resulting integrals, we have:

$$\mathbb{E}_{\mathbf{v}}\left[\sum_{i=1}^{n} p_i(\mathbf{v})\right] = \sum_{i=1}^{n} \mathbb{E}_{\mathbf{v}}\left[p_i(\mathbf{v})\right] = \sum_{i=1}^{n} \mathbb{E}_{\mathbf{v}}\left[\phi_i(v_i) \cdot x_i(\mathbf{v})\right] = \mathbb{E}_{\mathbf{v}}\left[\sum_{i=1}^{n} \phi_i(v_i) \cdot x_i(\mathbf{v})\right] \tag{3.14}$$

The RHS of formula (3.14) is the exact expected revenue of a mechanism with allocation rule $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$. If we remove all the ϕ's from the above equation, we are left with the expected social welfare of the auction. For this reason, this quantity is called the *expected virtual welfare* of an auction. For the ϕ's, formally we have:

Definition 3.16. ***(Virtual Valuation Function, [3]).*** *The virtual valuation $\phi_i(v_i)$ of bidder i with valuation v_i drawn from the distribution F_i is:*

$$\phi_i(v_i) = v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \tag{3.15}$$

An important remark is that the virtual valuation of an agent depends on her valuation and her distribution, but not on those of the other agents.
From formula (3.14) it becomes clear that if we want to maximize the expected revenue, we have to choose an allocation rule that maximizes the expected virtual social welfare. But we havenŷet made sure that the resulting mechanism is truthful. In order for that to be the case, from Myerson's lemma the allocation rule has to be *monotone*, meaning that if an agent bids more, her allocation can only increase. From formula (3.14) it follows immediately that an equivalent condition is that the virtual valuation function of every bidder is monotone non-decreasing with respect to her bid. Formally:

Definition 3.17. ***(Regular Distribution, [3]).*** *A distribution F is regular if the corresponding virtual valuation function $\phi_(v) = v - \frac{1-F(v)}{f(v)}$ is nondecreasing.*

Note: In some books, for a distribution to be called regular the corresponding virtual valuation function has to be strictly increasing. In most applications however, that assumption can be relaxed to allow non-decreasing virtual valuation functions.
An important subclass of regular distributions that arises often in practice is *monotone hazard rate* distributions. For these distributions, the second, negative term of the virtual valuation function is by itself monotone non-decreasing. Formally:

Definition 3.18. ***(Monotone Hazard Rate Distribution (MHR), [3]).*** *A distribution F satisfies the monotone hazard rate condition if $\frac{f(v)}{1-F(v)}$ is monotone non-decreasing in the support of the distribution.*

Pictorially, a MHR distribution has tails that decay at least as fast as those of an exponential distribution.
Now we are ready to define the mechanism that maximizes expected revenue in single-parameter environments. This is commonly referred to as *Virtual Surplus Maximization mechanism* (VSM) or simply Myerson mechanism. As we hinted to earlier, for the mechanism to be DSIC the valuation distributions of all bidders have to satisfy the regularity condition. Alternatively, they can be *ironed* to make them monotone, while at the same time preserving the virtual welfare of the auctions that matter. For a textbook treatment of the topic, we refer the interested reader to [3, Chapter 3].

Definition 3.19. ***(Myerson Mechanism, [10]).*** *Consider an arbitrary single-parameter environment and valuation distributions $F_1, F_2, \cdots F_n$. The virtual welfare-maximizing allocation rule is now defined as that which, for for each input $\boldsymbol{v}$ chooses the feasible allocation that maximizes the virtual welfare $\sum_{i=1}^{n} \phi_i(v_i)x_i(\boldsymbol{v})$. If every distribution F_i is regular, then this allocation rule is monotone, Coupling it with the unique payment rule of Myerson to meet the DSIC constraint, we obtain the revenue-optimal auction.*

Chapter 4

Frontiers of Mechanism Design

In this chapter we will overview some of the more modern approaches to mechanism design that have become prevalent in the last years.

4.1 PAC Learning and Automated Mechanism Design

As we alluded to in the previous chapter, the design of revenue-maximizing mechanisms is a notoriously challenging problem with tremendous real-world impact. Namely, even for 2 items available for sale and 2 bidders, the revenue maximizing mechanism is not known. Additionally, even for the settings where the revenue-maximizing mechanism is known, in order for it to work it requires knowledge of the distribution over the bidders' valuations. In an attempt to circumvent these two fundamental issues, the idea of using samples to guide the design of mechanisms was born. Sample-based mechanism design was first introduced in the context of *automated mechanism design*. In automated mechanism design, the goal is to design algorithms that take as input information about the set of bidders and return a mechanism that maximizes some objective function, such as revenue [32, 33, 34]. In the early days of automated mechanism design, the input information about the bidders was an explicit description of the distribution over their valuations. The support of this distribution is often doubly exponential, rendering obtaining and storing it prohibitive. In response, sample-based mechanism design was introduced where the input is now a set of samples from this distribution [35, 36, 37]. The mechanism designer's goal is to field a mechanism with high expected revenue on the distribution over the agents' valuation functions. One of the most common uses of samples in automated mechanism design is for the mechanism designer to pick some mechanism class that should theoretically "perform well" for the given setting, calculate the *empirical revenue* on those samples for different mechanisms of that class and then field the *empirically optimal mechanism*, i.e. the one that performed the best on the samples. Unfortunately, if the set of mechanisms that the mechanism designer optimizes over is complex, a mechanism may have high empirical revenue on a sufficiently small set of samples but low expected profit.

The solution to this problem is given by a framework of learning theory called *probably approximately correct learning* (PAC learning). In this framework, the learner receives some samples and is tasked with selecting a generalization function, called the *hypothesis*, from a certain class of possible functions. The learner's goal is that, with a high probability (the "probably" part), the selected function will have low "generalization error" (the approximately part), meaning that its expected performance in the real world will be close to its performance in the samples. But how can this framework be applied to automated mechanism design? Suppose the auctioneer is trying to field the optimal mechanism from some class of mechanisms $\mathcal{C}$. For any mechanism in that class, we can express its revenue on any valuation profile as a function only of $\boldsymbol{v}$, i.e. for any mechanism $\mathcal{M} \in \mathcal{C}$, $\text{Rev}(\mathcal{M}, \cdot)$ is a function only of $\boldsymbol{v}$. Then, instead of the auctioneer trying to maximize directly over the mechanism class $\mathcal{C}$, she can instead maximize over the set of their corresponding revenue functions, $\{\text{Rev}(\mathcal{M}, \cdot) \mid \mathcal{M} \in \mathcal{C}\}$ and simply field the mechanism with the best performing revenue function from that set. Because those revenue functions are, from a mathematical standpoint, simple scalar functions, the PAC learning framework can be applied to them and provide *statistical guarantees* about how close their performance in the real world will be to what it was in the samples. Those statistical guarantees are dependent on the number of samples used to guide the selection of the revenue function, and therefore the mechanism, and and some measure of the *intrinsic complexity* of the set of functions the mechanism designer maximized over. In this section we will overview those complexity measures and how they affect the statistical guarantees that PAC learning can provide.

4.1.1 Pseudo-Dimension and Connection to Learning Theory

Perhaps the first concrete effort in computational learning theory of explaining the learning process from a statistical point of view was developed by Vladimir Vapnik and Alexey Chervonenkis during 1960-1990, in the aptly named *Vapnik-Chorvenkis* theory. At the heart of this theory lies the definition of *VC dimension*.

Definition 4.1 (VC dimension). *A classification model f with some parameter vector $\boldsymbol{\theta}$ is said to shatter a set of data points $(x_1, x_2, \ldots, x_n)$ if, for all assignments of labels to those points, there exists a $\boldsymbol{\theta}$ such that the model f makes no errors when evaluating that set of data points. The VC dimension of a model f is the cardinality of the largest set that can be shattered by f.*

The keen reader may have noticed that the VC dimension definition referred to a classification problem. But in automated mechanism design the problem of revenue-maximization that the mechanism designer faces is real-valued. The real-valued analog of VC dimension was introduced by Pollard in 1984 and is called *pseudo-dimension*. Formally:

Definition 4.2 (Pseudo-Dimension, [5]). *The pseudo-dimension of a function class F is the cardinality of the largest set $S = \{x_1, x_2, \ldots, x_N\}$ and respective thresholds $y_1, y_2, \ldots, y_N$ such that all 2^N above/below patterns can be achieved by functions $f \in F$.*

Intuitively, the idea is to convert the real-valued functions back to a binary classification model, and use VC dimensions. To do that, we supplement the sample set with appropriately-chosen thresholds for each sample, and now the binary classification of each sample is whether the real-valued function is above or below the corresponding threshold.
The main theorem bridging the gap between *automated mechanism design* and *computational learning theory* is Pollard's uniform convergence theorem. If applied to mechanism design, this theorem provides a direct bound on the *generalization error* of a mechanism class, meaning for any mechanism in the class, it upper bounds how far off its empirical revenue on the samples can be from its actual expected revenue, in relation to the size of the sample set and the pseudo-dimension of the class. Formally:

Theorem 4.1 (Uniform Convergence, [12]). *For any $\delta \in (0,1)$ and any distribution $\mathcal{D}$ over v, with probability at least $1-\delta$ over the draw $\{v^1, v^2, \ldots v^N\} \sim \mathcal{D}^n$, for any mechanism $\mathcal{M} \in \mathcal{C}$,*

$$\left| \mathbb{E}_{v\sim\mathcal{D}}[Rev(\mathcal{M}, v)] - \frac{1}{N}\sum_{i=1}^{N} Rev(\mathcal{M}, v^i) \right| = O\left(U\sqrt{\frac{Pdim(\mathcal{C})}{N}} + U\sqrt{\frac{\log(1/\delta)}{N}} \right) \tag{4.1}$$

where U: The maximum revenue achievable by mechanisms in $\mathcal{C}$

4.1.2 Generalization and Representation Error Bounds

In the previous subsection we explored how the mechanism designer can utilize the sample complexity of a mechanism class to effectively prove that the class *generalizes well*, meaning that for any mechanism in the class, its empirical revenue in the real world is close to its empirical revenue in the samples. The keen reader may then wonder why the mechanism designer wouldn't then always optimize over some fixed mechanism class with very low pseudo-dimension, thus ensuring that the generalization error bound is very low. The answer is that in this case, the mechanism class may exhibit high representation error: Indeed for any mechanism in the class its performance in the real world is similar to its performance in the samples, but the class probably only contains mechanisms that perform poorly, both in the samples and in the real-world.
When trying to provide statistical guarantees for the performance of a mechanism class, another component equally important to the generalization error bound of the class is the *representation error bound*. Informally:

Definition 4.3. ***(Representation Error, [13]).*** *The representation error of a mechanism class $\mathcal{C}$ is the amount of revenue sacrificed, compared to the optimal mechanism, by restricting the search space to auctions in $\mathcal{C}$.*

The representation error depends on the setting in question each time, and in the distributions of the bidders valuations. This is the central challenge faced in automated mechanism design: The mechanism designer needs to identify, for the setting in question each time, a mechanism class $\mathcal{C}$ that balances representation error, i.e. the revenue sacrificed, with generalization error, i.e. the learning error incurred when learning auctions from that class. If the mechanism designer combines those 2 losses, she can then reason about how far-off the expected revenue of the mechanism that she chose is from the expected revenue of the optimal mechanism, in relation to the number of the samples used. Equivalently, the mechanism designer can calculate how many samples are required so that the empirically optimal mechanism on the samples has expected revenue that is a $(1-\epsilon)$-approximation of the optimal one. This is known as the *sample complexity* of the mechanism class.

4.1.3 Balcan's Sample Complexity Framework

One of the most impactful papers in automated mechanism design was the 2018 paper "A General Theory of Sample Complexity for Multi-Item Profit Maximization" by Maria-Florina Balcan et al. In this paper the writers explored in depth how the mechanism designer can use intrinsic properties of the structure of the mechanism class to prove strong sample complexity bounds. This process requires a sharp understanding of the interplay between mechanism parameters and buyer values, and how this interplay affects the final result of the mechanism. Then, they applied their theoretical results to many not well-understood mechanism classes, proving new sample complexity bounds or improving existing ones. In this thesis we will utilize their theoretical results in a similar fashion in order to prove sample complexity bounds for the mechanism classes that we introduce.
The first concept the writers introduced in this paper was that of *delineable mechanism classes*.

Definition 4.4 (((d,t)-delineable, [12]). *We say a mechanism class $\mathcal{C}$ is (d,t)-delineable if:*

1. *The class $\mathcal{C}$ consists of mechanisms parameterized by vectors $\boldsymbol{p}$ from a set $\mathcal{P} \subseteq \mathbb{R}^d$; and*

2. *For any $\boldsymbol{v}$ in the support of the distribution D over buyers' valuations, there is a set $\mathcal{H}$ of t hyperplanes such that for any connected component $\mathcal{P}'$ of $\mathcal{P} \setminus \mathcal{H}$ the function $Rev(\boldsymbol{p}, \boldsymbol{v})$ is linear over $\mathcal{P}'$.*

This is a novel way of viewing the intrinsic complexity of the structure of a mechanism class: Fix any valuation profile $\boldsymbol{v}$ in the support of the distribution of the bidders' valuations. The revenue on that fixed valuation profile can be viewed as a function of only the parameter vector $\boldsymbol{p}$ of the mechanism in that class. Then, the parameter space can be broken into subregions where the profit on that valuation profile is linear in $\boldsymbol{p}$. Intuitively, the higher the number of those subregions, the higher the expressive capabilities, and therefore complexity, of the mechanism class. Balcan et al. formalized this intuition in the following theorem:

Theorem 4.2. *If $\mathcal{C}$ is (d,t)-delineable, the pseudo-dimension of $\mathcal{C}$ is $O(d \log(dt))$.*

4.2 Deferred-Acceptance Auctions

4.2.1 Preliminaries

The framework of Deferred Acceptance (DA) auctions was introduced by the work of Milgrom and Segal ([14]). This family of mechanisms was initially restricted to single-minded bidders: Let N be the set of bidders and B_i the *bid space of bidder i*, i.e. the total range of possible values for her bid. In the auction, each bidder can either "win", which means that she is serviced, or "lose" which means that she is not serviced and pays nothing. Every bidder has zero value for losing, and some value for winning, which is what she declares to the mechanism. The preferences of each bidder only depend on whether she wins or loses, and how much she has to pay in case she wins, i.e. they don't depend on the set of the other winning bidders. Informally, in their first introduction, deferred-acceptance auctions were a family of adaptive, backwards-greedy algorithms for any such single-minded setting.

Definition 4.5 (Deferred-Acceptance Auction, [15])**.** *A deferred-acceptance auction is a particular kind of mechanism described by a set of scoring functions, as follows. The set of bidders that have not yet been finalized by the mechanism are called active. For each set $A \subseteq N$ of active bidders and each bidder $i \in A$, there is a scoring function $s_i^A : B_i \times B_{N \setminus A} \to \mathbb{R}_+$ that is nondecreasing in its first argument. The auction then operates as follows. Let $A_t \subseteq N$ denote the set of active bidders in stage t. Initialize $A_1 = N$. For each $t \geq 1$, if the set of active bidders constitute a feasible solution, then stop the auction and output A_t as the set of winning bidders, otherwise, $A_{t+1} = A_t \setminus \arg\min_{i \in A_t} s_i^{A_t}(b_i, b_{N \setminus A_t})$ and continue.*

Intuitively, at each round the deferred-acceptance auction is *finalizing*, i.e. removing from the active set, the bidder that seems the least promising, according to her scoring function.

Up to this point we have not defined the payment rule of a deferred-acceptance auction. In a sense, we don't have to: The environment is single-parameter and the scoring functions monotone non-decreasing in the bids of the respective players. Thus, for the auction to be DSIC, the payments have to be the ones specified by Myerson's Lemma. However, due to the special structure of deferred-acceptance auctions, there exists an elegant algorithm for calculating the payments of the winning bidders:

Algorithm 2: Calculating Payments in DA Auctions

Result: For each winning bidder $i \in A_T$, $p_i^T(b)$ is her payment.

$p_i^0 \leftarrow \inf B_i \;\; \forall$bidder i;
for *each round* $t \geq 1$ **do**
  for *each winning bidder* i **do**
    $p_i^t(b) \leftarrow \max\{p_i^{t-1}, \inf\{b_i' \in B_i : s_i^{A_t}(b_i', b_{N\setminus A_t}) > s_j^{A_t}(b_j, b_{N\setminus A_t})$ for $j \in A_t \setminus A_{t+1}\}\}$
  end
end

In words, each winning bidder pays the minimum she could have been and still not gotten finalized in any round of the auction, holding the bids of the other bidders fixed. A more careful inspection of the algorithm reveals a very critical property of deferred-acceptance auctions: holding fixed the final set of winners A_T, winning bidders' payments depend only on the bids of the losing bidders, $b_{N\setminus A_t}$, and not on the bids of the winning bidders b_{A_T}. This implies that no winning bidder can affect another winner's payment, except by changing to a losing bid. In this chapter we will explore in detail the importance of this property.

4.2.2 Clock Auctions and Equivalence to Deferred-Acceptance Auctions

Most of the remarkable properties of deferred-acceptance auctions stem from the fact that they are equivalent to *clock auctions*. Thus, in order for one to better understand those properties, a brief introduction to clock auctions is tantamount. Informally, an (ascending) clock auction proposes an increasing sequence of prices to each bidder, with each new offer followed by a decision period, in which that bidder whose price was strictly increased can decide whether she wants to exit or to continue in the auction. Bidders that have never exited are called *active*, and those that have are called *inactive*. Bidders who decide to continue when their price is increased are said to *accept* the new offer. When the mechanism terminates, the still-active bidders are the winners of the auction, and they each pay their highest accepted price. The differentiating factor between different clock auctions is their pricing functions, which determine the prices offered to each bidder. Formally:

Definition 4.6 (Ascending Clock Auction, [15])**.** *A period-t history consists of the sets of active bidders in all periods up to period t, i.e.* $A^t = (A_1, A_2, \ldots, A_t) \in (2^N)^t$ *such that* $A_t \subseteq A_{t-1} \subseteq \cdots \subseteq A_1$. *Let H denote the set of all such histories. An ascending clock auction is defined by a price mapping* $p : H \to \mathbb{R}^N$ *such that for all* $t \geq 2$ *and all* A^t, $p(A^t) \geq p(A^{t-1})$. *The clock auction initializes* $A_1 = N$. *In each period* $t \geq 1$, *given history* A^t, *it offers prices* $p(A^t)$ *to bidders. If* $p(A^t) = p(A^{t-1})$, *the auction stops and bidder i is a winner if and only if* $i \in A^t$, *and in that case she pays* $p_i(A^t)$. *If* $p(A^t) \neq p(A^{t-1})$, *then each bidder in* A^t *chooses whether or not to exit the set of active bidders. Letting* $E \subseteq A_t$ *denote the set of bidders who chose to exit, the auction continues in period* $t+1$ *with the new set of active bidders* $A_{t+1} = A_t \setminus E_t$ *and new history* $A^{t+1} = (A^t, A_{t+1})$.
To complete the description of the auction as an extensive-form mechanism, we also need to describe bidders' information sets. General information disclosure is allowed in clock auctions: bidder i observes some signal $\sigma_i(A^t)$ *in addition to her current price* $p_i(A^t)$ *in history* A^t.

Even though the formal description of clock auctions is quite involved, the strategies that bidders implement are, in most cases, simple. Intuitively, every bidder accepts the offer of the mechanism as long as it is below her threshold, which in this case is called her *cutoff*. Formally:

Definition 4.7 (Cutoff Strategy, [15])**.** *A strategy for bidder i in a clock auction is a cutoff strategy with cutoff b_i if it specifies exit if and only if $p_i(A^t) > b_i$, for some $b_i \geq p_i(N)$.*

The reasoning behind the condition $b_i \geq p_i(N)$ is that every bidder accepts her opening offer, as it is usually some arbitrarily small value.

Definition 4.8 (Finite Clock Auction)**.** *We say that a clock auction is finite if there exists some T such that the auction stops by period T.*

Example 4.1. *Let's say we have* 3 *cinema tickets available for sale, and* 5 *bidders that all want to go to the cinema. Bidder's one through five value for watching the movie is* $3, 4, 5, 6,$ *and* 7 *respectively. An ascending clock auction we could implement in this scenario is the following: Every bidder is contesting one ticket. Initially, we offer a price of* 0 *to every bidder. As long as the bidders that accept the current offer are more than* 3 *(the number of available tickets), we raise the asking price by 1 and continue. When the active bidders that remain are less or equal to* 3 *we stop, and each bidder pays the latest accepted offer. Then, the cutoff strategies of the bidders are also very simple: Every bidder accepts the offer of the auction, as long as that offer is less or equal to her value for watching the movie. In our example, bidder* 1 *will be the first that declines the asking price, when the price becomes* 4*, and bidder* 2 *will exit second, one period later, when the asking price becomes* 5*. The 3 remaining bidders will all accept the asking price of* 5*, the auction will terminate, and each of them will pay a price of* 5 *for their ticket.*

As mentioned earlier, the reason for discussing clock auctions in this thesis is their equivalence to deferred-acceptance auctions, and the fact that this is the source of some of the remarkable properties of the latter. This equivalence was first highlighted by Milgrom and Segal when they introduced the framework of deferred-acceptance auctions. The following 2 propositions show exactly this equivalence.

Proposition 4.1 ([15])**.** *For every deferred-acceptance auction with finite bid spaces and threshold pricing, there exists an equivalent clock auction in which bidders are restricted to cutoff strategies.*

Proof. Given bid spaces $B_1, B_2, \ldots, B_N$, for each $v \in \mathbb{R}$ and bidder i let $v^+ = \min\{b_i \in B_i : b_i > v\}$ and $v^- = \max\{b_i \in B_i : b_i < v\}$. Let the opening prices be $p_i(N) = \min B_i$ for each bidder i. Given a deferred-acceptance auction with scoring rule s, we can construct an equivalent clock auction as follows: The price increase rule in the clock auction increases the price offered to each lowest-scoring bidder by the minimal amount, while leaving the prices unchanged for the other bidders:

$$p_i(A^t) = \begin{cases} p_i(A^{t-1})^+, & \text{if } i \in \arg\min_{j \in A_t} s_j^{A_t}\left(p_j(A^{t-1}), p_{N \setminus A_t}(A^t)^+\right) \\ p_i(A^{t-1}) & \text{else.} \end{cases} \tag{4.2}$$

An important remark is that the clock auction maintains $p_i(A^t) = p_i(A^{t-1})\ \forall i \in N \setminus A_t$, effectively memorizing the prices rejected by the bidders who have quit the auction. Thus, their cutoffs can be inferred as $p_i(A^t)^-$.

The equivalence becomes apparent: First, for every history of the clock auction, the next set of bidders to quit in the clock auction is the set of bidders who have the lowest score among the set of active bidders, thus the set of winners is the same in both auctions. Additionally, if any winning bidder had declined any lower price, she would have exited at that point, so each bidder's final clock price is the lowest cutoff she could have used and still won. □

Proposition 4.2 ([15])**.** *For every finite clock auction in which bidders are restricted to cutoff strategies, there exists an equivalent deferred-acceptance auction with finite bid spaces and threshold prices.*

Proof. Given a finite clock auction P, we can construct bid spaces and and a scoring function to create an equivalent deferred-acceptance auction. We take each bidder i's space to be $B_i = p_i(h) : h \in H$, i.e. the set of possible prices agent i could face in the clock auction.
Next, we construct the scoring function in the following manner: Holding fixed a set of bidders $S \subseteq N$, and their bids $b_S \subseteq N^S$, let $A_t(S, b_S)$ denote the set of active bidders in the clock auction at round t in which every bidder $j \in S$ uses cutoff strategy b_j and every bidder from $N \setminus S$ never exits. Formally, initialize $A_1(S, b_S) = N$ and iterate by setting

$$A_{t+1}(S, b_S) = A_t(S, b_S) \setminus \{j \in S : b_j < p_j(A^t(S, b_S))\}. \tag{4.3}$$

This gives an infinite sequence $\{A_t(S, b_S)\}_{t=1}^{\infty}$ but the sets start repeating at the point where the clock auction stops.
Now for given $A, b_{N\setminus A}, i \in A$ and b_i, define the score of agent i proportional to how long she would remain active in the clock auction if she uses cutoff strategy b_i and all bidders in $N \setminus A$ use cutoffs $b_{N\setminus A}$, while bidders in $A \setminus \{i\}$ never quit:

$$s_i^A(b_i, b_{N\setminus A}) = \sup\{t \geq 1 : i \in A_t\left(\{i\} \cup (N \setminus A), (b_i, b_{N\setminus A})\right)\}. \tag{4.4}$$

(Note that the score is ∞ in the case that the auction stops with agent i still active.) This score is by construction non-decreasing in b_i. Also by construction, given a set A of active bidders, the set of bidders to be rejected by the scoring function in the next round ($\arg\min_{i\in A} s_i^A(b_i, b_{N\setminus A})$) is the set of bidders who would quit the soonest in the clock auction given the inactive bidders have used cutoffs $b_{N\setminus A}$. If no more bidders would exit the auction, then all active bidders have a score of infinity, so the auction stops. Finally, as argued above, the winners' clock auction prices are their threshold prices: the winner would have lost by using any lower cutoff in B_i than her clock auction price. □

4.2.3 Incentive Properties and Use in Practice

As we alluded to earlier, the motivation behind deferred-acceptance auctions is their many remarkable incentive properties. First of all, as we showed in the previous section, DA auctions are implementable as ascending (or descending, in the case of procurement) clock auctions. This in turn has many remarkable consequences. The primary one is perhaps the fact that clock auctions, and by extension DA auctions when implemented as clock auctions, are *obviously strategyproof*, in the sense formalized by Li:

Definition 4.9 (Obviously Dominant Strategy, [16])**.** *A strategy is obviously dominant if, for any deviation, at any information set where both strategies first diverge, the best outcome under the deviation is no better than the worst outcome under the dominant strategy.*

Definition 4.10 (Obviously Strategyproof Mechanism)**.** *A mechanism is obviously strategyproof (OSP) if it has an equilibrium in obviously dominant strategies.*

Despite the perhaps daunting formal definition, this has a very sound behavioral interpretation: A strategy is obviously dominant if and only if even a *cognitively limited* agent can recognize it as weakly dominant.
In practice, this is oftentimes a very desirable property: Let's take as an example the FCC Incentive Auction for reallocating spectrum. In simple terms, the aim of this auction was to

take back spectrum licenses that weren't used very efficiently, perhaps because they had been allocated a long time ago, to small, local TV stations, and allocate them to someone else who could put them to better use, for example a mobile network operator. It is apparent that the participation in the reverse auction part of those small TV stations was critical. However, those small TV stations didn't have high budgets to hire expensive consultants to advise them on how to bid on a complex mechanism. For their participation in the mechanism, obvious strategyproofness of the mechanism was necessary. In contrast even mechanisms that seem trivial to a mechanism designer, such as the sealed-bid Vickrey auction, are not obviously strategy-proof in this sense.

An additional notable property of DA auctions is that they are weakly group-strategyproof (WGSP). This means that no coalition of bidders can collude against the mechanism designer and collectively submit false bids in such a way that all members of the coalition are strictly better off. Effectively, no bidder has an incentive to collude against the mechanism designer. In practice, this is a very important property: In the modern world, assuming that agents in a game cannot or will not communicate, especially in high-impact markets with few agents, might be unrealistic. For those reasons, the reverse auction part of the recently concluded FCC incentive auction for reallocating spectrum was a DA auction [17]. For further advantages and motivation behind DA auctions we refer the interested reader to [14, 15].

4.2.4 Extension to Multiple Levels of Service

Prior to the work of Gkatzelis, Markakis and Roughgarden, all work on DA auctions was restricted to binary, single-parameter environments, where each bidder could either "win" or "lose" and had some value for winning. However, in the real world things are oftentimes more complicated than that: there exist multiple "*levels of service*", as opposed to simply a binary decision. Let's take as an example a train ticket purchase. Depending on the buyer's willingness to pay, the available choices may include third, second and first class. In a multi-unit auction with identical goods, the levels of service correspond to how many units of the good awarded to the bidder. Gkatzelis, Markakis and Roughgarden [18] extended the framework of DA auctions to such non-binary settings, and showed how it could be applied to some basic mechanism design problems.

Initially, they extended the framework of DA auctions to non- binary, single-parameter settings: There are multiple levels of service that each bidder can receive. Every bidder declares her value v per level of service. Then, if she receives some level of service l, her value for that level of service is $v \cdot l$. In this thesis we also refer to this environment as *linear bidders*, because the bidders' valuations are linear in their level of service.

The key idea in order to extend the framework of DA auctions to non-binary settings was to supply each auction with a *clinching function*, a special form of an allocation function. For every auction, this function dictates the level of service that every bidder is allocated, based on the number of rounds that she has remained active in the auction, her bid, and the bidders that have already exited from the auction. Just like the DA auctions introduced by Milgrom and Segal [14], these auctions operate in a sequence of stages. In the first stage all bidders are active and after each stage, the bidder with the lowest score is finalized according to her clinching function and she exits the mechanism. Finally, prices are simply set according to Myerson's Lemma. In order to retain the remarkable properties of the binary, single-parameter DA auctions introduced by Milgrom and Segal, careful consideration of the scoring and clinching functions is required. Formally:

Definition 4.11 (Generalized Single-Parameter DA Auction, [18])**.** *A generalized single-parameter auction operates in discrete stages $t \geq 1$. We denote by $A_t \subseteq N$ the set of currently* *active bidders in the beginning of stage t; initially, $A_1 = N$, and $A_{t+1} \subset A_t$ for every $t \geq 1$. The DA auction is fully defined by 2 sets of functions:*

- *The **scoring functions** $\sigma_i^{A_t}(b_i, \boldsymbol{b}_{N \setminus A_t})$, that are non-decreasing in their first argument.*
- *The **clinching functions** $g_i^{A_t}(\boldsymbol{b}_{N \setminus A_t})$, which are non-increasing with respect to the set of active bidders, i.e. $g_i^{A_{t+1}}(\boldsymbol{b}_{N \setminus A_{t+1}}) \geq g_i^{A_t}(\boldsymbol{b}_{N \setminus A_t})$.*

At each stage t, if $A_t \neq \emptyset$, then the level of service of some active bidder $i \in \arg\min_{i \in A_t}\{\sigma_i^{A_t}(b_i, \boldsymbol{b}_{N \setminus A_t})\}$ is finalized, possibly with the use of some tie-breaking rule. That is, a bidder i with the lowest score stops being active, we set $A_{t+1} = A_t \setminus \{i\}$, and her level of service is finalized at level $g_i^{A_t}(\boldsymbol{b}_{N \setminus A_t})$. When we reach $A_t = \emptyset$, then the auction terminates and the payment of each bidder is determined by Myerson's Lemma.

There are some worthwhile remarks to be made: The conditions that the scoring function of every bidder is non-decreasing in the bid of that player and that her clinching function is non-increasing with respect to the set of active bidders together ensure that the *allocation rule* of the mechanism is monotone non-decreasing with respect to the bid of the bidder: From Myerson's lemma, for any single-parameter environment, in order for any mechanism to be truthful, its allocation rule must be monotone. Another important remark is that neither the scoring nor the clinching function of any bidder depends on the bids of the other still active players. This condition ensures the weak group-strategyproofness of all generalized DA auctions.
Then, Gkatzelis, Markakis and Roughgarden further generalized the framework of DA auctions to multi-parameter environments with *submodular bidders*, i.e. bidders whose value for an additional level of service weakly decreases with the level that they have already clinched. Let $L = \{1, 2, \ldots, k\}$ be the set of possible levels of service. The difference compared to the case of budget-additive bidders is that now every bidder doesn't report a single value, her value per level, but instead reports k values, her *marginal values* for each subsequent level of service. Once again, in order to maintain the incentive properties of DA auctions, careful consideration of the scoring and clinching functions is required. Formally:

Definition 4.12 (Multi-Parameter Generalized DA auction, [18]). *A multi-parameter DA auction operates in stages $t \geq 1$. In each stage t a set of bidders $A_t \subseteq N$ is active. Initially $A_1 = N$ and $A_{t+1} \subset A_t$ for every $t \geq 1$. The DA auction is fully defined by two collections of functions:*

- *The **scoring functions** $\sigma_i^{A_t}(\cdot, \boldsymbol{b}_{N \setminus A_t})$ that are non-decreasing in their first argument.*
- *The **clinching functions** $g_i^{A_t}(\boldsymbol{b}_{N \setminus A_t})$, that are non-decreasing with respect to the set of active bidders, i.e. $g_i^{A_{t+1}}(\boldsymbol{b}_{N \setminus A_{t+1}}) \geq g_i^{A_t}(\boldsymbol{b}_{N \setminus A_t})$.*

At each stage t, if $A_t \neq \emptyset$, the score of bidder i is computed as:

$$\sigma_i^{A_t}(b_i(g_{it} + 1), b_{N \setminus A_t}), \tag{4.5}$$

i.e. the score is a function of the bidder's marginal value for receiving a level increase, given the level that she has already clinched. At every stage, the bidder with the lowest score is finalized at level $g_{it} = g_i^{A_t}(\boldsymbol{b}_{N \setminus A_t})$ and she is removed from the set of active bidders ($A_{t+1} = A_t \setminus \{i\}$).

If defined in this way, all the incentive guarantees of single-parameter DA auctions carry over in this setting.

Proposition 4.3 ([18]). *Every multi-parameter DA auction is weakly group-strategyproof, and has an equivalent clock auction implementation that is obviously strategyproof.*

4.2.5 Performance and Limitations of Deferred Acceptance Auctions

Ever since their inception, the performance of deferred-acceptance auctions has been studied extensively in many different settings and from different viewpoints. Milgrom and Segal [14] initially introduced procurement DA auctions, motivated by FCC auctions for reallocating spectrum. In a *procurement auction*, the bidders are the ones that possess the goods for sale, and it is the auctioneer who is interested in buying. In a more recent version of that same paper, they studied the social welfare that can be attained by DA procurement auctions for near-matroid environments: In these settings, DA auctions can achieve near-optimal social welfare. This result intuitively makes sense: For matroid environments, there are both forward and backward greedy algorithms that are optimal [19].

The work of Milgrom and Segal was followed up by Dutting et al in [20], in which they explored the power and limitations of DA auctions from the viewpoint of approximation algorithms. On the positive side, they proved that for the object of social welfare, for single-minded bidders DA auctions can have an almost matching approximation ratio to the state of the art mechanisms in the literature, effectively strengthening the incentive guarantees of the known approximations. On the negative side, for knapsack auctions they showed that no DA auction can achieve an approximation ratio to the optimal social welfare that is sublogarithmic in the number of items. On the other hand, there are known WGSP mechanisms that achieve a constant approximation ratio in this setting.

Group-strategyproof mechanisms have also been studied independently of DA auctions, with a big portion of those papers focusing on cost-sharing mechanisms, e.g. [21, 22]. Most of this work was in the context of cost-sharing mechanisms. As is often the case, it was observed that the stronger incentive guarantees come at a significant cost in terms of efficiency.

Finally, Gkatzelis, Markakis and Roughgarden studied the performance of multi-parameter DA auctions in the setting of multi-unit auctions with submodular valuations [18]. They proved that in these settings, the VCG mechanism is not implementable as a DA auction, establishing a first gap between DA auctions and optimal mechanisms. They then strengthened it by proving that even for 2 players and 2 units, no WGSP mechanism can guarantee an approximation factor to the social welfare that is better than $\sqrt{2}$.

Most of the work thus far on the performance of deferred-acceptance auctions focused on the objective of social welfare maximization, and from the perspective of a *worst-case analysis.* In this thesis, we took a different approach on both of these critical points: We are attempting to determine settings in which, if we make some natural assumptions about the bidders' valuation distributions, we can design DA auctions that, given some samples of the bidders' distributions, achieve expected revenue close to optimal.

Chapter 5

Learning Revenue-Optimal Deferred-Acceptance Auctions for Single-Parameter Environments

In the following two chapters we will determine settings in which, given some samples of the bidders' valuation distributions, one can learn DA auctions with high expected revenue. In this chapter we will focus on single-parameter environments, specifically multi-unit auctions where each bidder only declares a single number in the auction, her value per unit of the good. For those environments we will determine sufficient conditions to learn DA auctions that achieve expected revenue arbitrarily close to optimal, and upper bound the sample complexity to do so. Then, we will extend those results to environments with polymatroid constraints.

5.1 Multi-Unit Auctions

5.1.1 t-level Auctions

For single-parameter environments, the mechanism classes we propose are modified versions of the t-level auctions introduced by Morgenstern and Roughgarden in [23] that we show can be implemented as DA auctions. For one to better understand the proposed mechanisms, a basic understanding of t-level auctions is necessary.
In [23], Morgenstern and Roughgarden focused exclusively on single-parameter environments. For those environments, they introduced the novel mechanism class of t-level auctions as a way of balancing the competing demands of expressivity and simplicity. First, they introduced the mechanism class for the singe-item case:

Definition 5.1 (Single-Item t-level Auctions). *For each bidder i there are t numbers $0 \leq l_{i,0} \leq l_{i,1} \leq l_{i,t-1}$. This set of tn numbers defines a t-level auction with the following allocation and payment rules. Consider a valuation profile $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$:*

1. *For each bidder i, let $t_i(v_i)$ denote the index τ of the largest threshold $l_{i,\tau}$ that lower bounds v_i (or -1, if $v_i < l_{i,0}$). We call $t_i(v_i)$ the level of bidder i.*

2. *Sort the bidders from highest level to lowest and, withing a level, use a fixed lexicographical tie-breaking ordering $\prec$ to pick the winner.*
3. *Award the item to the first bidder in this sorted order (unless $t_i = -1$ for every bidder i, in which case there is no sale).*

The payment rule is the unique one that render truthful bidding a dominant strategy and charges 0 to losing bidders: The winning bidder pays the lowest bid at which she would continue to win.

Intuitively, in a t-level auction there are t possible prices that a bidder i could face. Which of those she will actually face in the auction depends solely on the bids of the other bidders.

Example 5.1. *Consider the following 5-level auction for bidders a,b and c. Let* $l_{a,\cdot} = [2, 4, 6, 8, 10]$, $l_{b,\cdot} = [1, 4, 8, 10, 12]$ *and* $l_{c,\cdot} = [1.4, 4.3, 7, 11, 13]$. *For example, if bidder a bids less than 2 she is at level* -1*, a bid in* $[2, 4)$ *places her at level* 0*, a bid in* $[4, 6)$ *at level* 1*, a bid in* $[6, 8)$ *at level* 2*, a bid in* $[8, 10)$ *at level* 3 *and finally a bid of at least* 10 *places her at level* 4*. Let* $c \prec b \prec a$*. There are 3 distinct cases:*

- ***Monopoly prices:*** *If* $v_a < 2$, $v_b < 1$ *and* $v_c \in [4.3, 7)$*, then bidders a and b are at level* -1 *and bidder c is at level 2. So, bidder c wins and pays* 1.4*, the minimum she needs to be in order to bid at least at level* 0*.*
- ***Multiple at the highest level:*** *If* $v_a \in [4, 6)$, $v_b \in [4, 8)$ *and* $v_c \in [1.4, 4.3)$ *then bidders a and b are at level 2 and bidder c is at level 1. The highest level is 2, and there are 2 bidders at that level, a and b. Between them, the tie-breaking rule favors bidder a, so she wins and pays 4, the minimum she needs to bid to be in level* 2*.*
- ***Unique at the highest level:*** *If* $v_a > 10$, $v_b \in [4, 8)$ *and* $v_c \in [4.3, 7)$ *then bidder a is at level* 4*, and bidders b and c are at level* 1*. Bidder a wins because she is at the highest level, and she pays the* 4*, because even if she was not the only one at level* 1*, she would still win.*

The main theorem that Morgenstern and Roughgarden used to prove their sample complexity bounds for this mechanism class was Pollard's uniform convergence theorem. If solved for the number of samples, it can be stated as:

Theorem 5.1 ([5])**.** *Suppose* $\mathcal{C}$ *is a class of real-valued function with range in* $[0, H]$ *and pseudo-dimension* $Pdim(\mathcal{C})$*. Then, for every* $\epsilon > 0$ *and* $\delta \in [0, 1]$*, the sample complexity of* (ϵ, δ)*-uniformly learning f with respect to* $\mathcal{C}$ *is* $N = O\left(\left(\frac{H}{\epsilon}\right)^2 \left(Pdim(\mathcal{C}) \ln\left(\frac{H}{\epsilon}\right) + \ln\frac{1}{\delta}\right)\right)$

One important remark is that this guarantee is realized by the learning algorithm that simply outputs the function (or mechanism, in the case of mechanism design) with the smallest empirical error on the sample set. This is called empirical risk minimization (ERM).
Using this family of mechanisms, they managed to show that only a polynomial number of samples is required to achieve expected revenue arbitrarily close to optimal in the single-item setting. Their proof involved 2 main parts. The first part was showing that auctions of this kind have small *representation error*, meaning that for every product distribution F over bidders' valuations, this family of mechanisms contains an auction with relatively small t and expected revenue close to optimal. The second part was to show that this mechanism class has low pseudo-dimension, and therefore leads to small generalization error. We will defer those proofs as we will need to perform relatively the same process later in our thesis.
They also extended their results to *matroid environments*. The generalization is pretty straightforward: They order the bidders by level, breaking ties within a level by the fixed lexicographical tie-breaking ordering, $\prec$, and greedily choose winners according tho this ordering, subject

to feasibility constraints. Matroid environments have been well-studied in mechanism design. Formally, they can be defined as:

Definition 5.2 (Matroid Environment)**.** *A collection $\mathcal{X}$ of subsets is a matroid iff it satisfies 2 properties:*

1. *Whenever $X \in \mathcal{X}$ and $Y \subseteq X$, $Y \in \mathcal{X}$ and*
2. *For two sets $|I_1| < |I_2|$, $I_1, I_2 \in \mathcal{X}$, there is always an augmenting element $i_2 \in I_2 \setminus I_1$ such that $I_1 \cup \{i_2\} \in \mathcal{X}$.*

5.1.2 Warm-Up: Unit-Demand Bidders

The first setting we will study is a *multi-unit auction with unit demand bidders*, i.e. there are m identical copies of the good available for sale, and each bidder is interested in purchasing one unit. One such example would be a selling bus tickets, where the bus only has a specific number of seats available. This environment is a special case of a multi-unit auction with bidder-specific demands that we will study later. The reason for exploring it separately is to smoothly introduced the reader to an application of t-level auctions, and how they can be implemented as DA auctions. First, we will have to formally define the setting.

The Setting (unit-demand bidders)

- n bidders.
- multi-unit auction with m units available for sale.
- *unit-demand bidders:* Each bidder is interested in purchasing one unit of the good.
- each bidder i's value per unit, v_i, follows a distribution with probability density function $f_i(\cdot)$ and virtual valuation function $\phi_i(\cdot)$.
- The valuation distributions of all the bidders are bounded in $[1, H]$.

The mechanism class we propose for this setting is in essence the t-level matroid auctions introduced in [23]. Our main contribution in this subsection is showing how those t-level auctions can be implemented as DA Auctions.

Definition 5.3 (Unit-Demand DA t-level Auctions)**.** *Each bidder i faces t thresholds: $0 \leq l_{i,0} \leq l_{i,1} \leq \cdots \leq l_{i,t-1}$. This set of numbers defines a Unit-Demand DA t-level auction with the following allocation rule. Consider a valuation profile $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$:*

1. *For each bidder i, let $t_i(v_i)$ denote the index τ of the largest threshold $l_{i,\tau}$ that lower bounds u_i (or -1 if $v_i < l_{i,0}$). We call $t_i(v_i)$ the level of bidder i.*
2. *Sort the bidders from lowest level to highest level and, within a level, use a fixed lexicographical tie-breaking ordering $\prec$ to pick a winner.*
3. ***DA Allocation:** Start finalizing the bidders, from lowest level to highest level. The first $n - m$ bidders to be finalized aren't allocated a unit. Every one of the final n bidders is allocated a unit when he is finalized, but only if his level is greater than -1.*

4. ***Payment Rule:*** *In accordance with t-level matroid auctions, holding the bids of the other bidders fixed, every bidder pays the minimum she could've bid and still won a unit.*

Note: The t notation is overloaded: t can denote either some level of a t-level auction or a specific round of a deferred-acceptance auction, depending on the context.

For this mechanism class, we will prove 2 things:

1. That any auction of the above mechanism class is a valid generalized DA auction, as defined in [18].
2. That any auction in the above class is also a valid t-level matroid auction as defined in [23]. This implies that we can use directly the sample complexity bounds from that paper.

Proposition 5.1. *Any unit-demand DA t-level auction, as described in definition 5.3, is a valid generalized DA auction.*

Proof. Let $0 \leq l_{i,0} \leq l_{i,1} \leq \cdots \leq l_{i,t-1}$ be the thresholds that bidder i faces in the unit-demand DA t-level auction. These tn numbers uniquely define a generalized DA t-level auction with the following set of scoring and clinching functions:

- scoring functions: $\sigma_i^{A_t}(b_i, \boldsymbol{b}_{N \setminus A_t}) = t_i(v_i)$
- clinching functions: $g_i^{A_t}(\boldsymbol{b}_{N \setminus A_t}) = \begin{cases} 0, & \text{if } |A_t| > m \\ 1, & \text{else} \end{cases}$

The scoring function of every bidder is weakly increasing in her bid since the thresholds she faces are weakly increasing and her level is simply the index of the highest level that is still lower than her bid. Also, the scoring function of every bidder does not depend on the bids of the other still active bidders.
The clinching function of every bidder does not depend on the bids of the other still active bidders and is obviously non-increasing with respect to the set of active bidders. Both the scoring and clinching functions satisfy all the conditions of definition 4.11, therefore any such auction is a valid generalized DA auction. □

Proposition 5.2. *Any unit-demand DA t-level auction, as described in definition 5.3 is a valid t-level matroid auction.*

Proof. First of all, the setting is a matroid:

- Suppose X is a feasible set of winners in the auction. Equivalently, X contains at most m winners. Then, for any $Y \subseteq X$, it holds: $Y \subseteq X \implies |Y| \leq |X| \leq m$. Thus, Y is also feasible.
- Suppose I_1, I_2 are 2 feasible sets of winners such that $|I_1| \leq |I_2|$. Then, I_2 necessarily contains at least 1 winner, i_2 that I_1 does not, else it would be $|I_1| \geq |I_2|$, a contradiction. Also, we have $|I_1| < |I_2| \leq m$, since I_2 is feasible. It immediately follows that $I_1 \cup i_2$ is feasible.

Now that we showed that the setting is a matroid, it suffices to prove that the allocation rule coincides with this of matroid t-level auctions. Then, because both mechanisms use Myerson payments, the payment rule will coincide also. But this is easy to see: For the exact same thresholds, both the DA auction described and the matroid t-level auction will result in the exact same allocation; they will allocate the m units to the m bidders with the highest levels, breaking ties in the same lexicographical ordering $\prec$, provided that their levels are non-negative. $\square$

From the previous 2 propositions, it follows immediately that in this setting, matroid t-level auctions are equivalent to the unit-demand DA t-level auctions of definition 5.3. Therefore, the sample complexity result of [23] for this mechanism class carries over:

Corollary 5.1 (Application of [23], Corollary 5.6)**.** *With probability $1-\delta$, the empirical revenue maximizer for a sample of size N of the class of Unit-Demand DA t-level Auctions is a $(1-O(\epsilon))$-approximation to Myerson for n bidders whose valuations are in $[1, H]$, for $t = O\left(\frac{1}{\epsilon} + \log_{1+\epsilon} H\right)$ and $N = O\left(\left(\frac{Hm}{\epsilon}\right)^2 \left(nt\log(nt)\ln\frac{Hn}{\epsilon} + \ln\frac{1}{\delta}\right)\right) = \widetilde{O}\left(\frac{H^2m^2n}{\epsilon^3}\right)$.*

This result was expected: In this environment, forward and backward greedy algorithms coincide. Therefore, matroid t-level auctions, that use a greedy algorithm, are also implementable as DA auctions.

5.1.3 From Unit-Demand to Budget-Additive Bidders

In this section we will extend the previous result to the setting of bidders with linear valuations: Now every bidder has an upper limit on the number of units she wants, named her *demand*, that is known to the auctioneer prior to the auction. In the auction every bidder once again only declares a single number: her value per unit of the good. This setting can capture from unit-demand up to unbounded-demand bidders. Formally:

The Setting (Budget-Additive Bidders)

- n bidders.
- multi-unit auction with m units available for sale.
- every bidder i has a publicly known demand d_i.
- every bidder is additive, up to her demand d_i:
  For every bidder i, her value for acquiring x_i units of the good (a service level x_i), up to d_i, is $x_i \cdot v_i$, where v_i: her value per unit of the good. For more than d_i units: her value remains $v_i \cdot d_i$. These valuation functions are called *budget-additive*.
- Every bidder i's value per unit, v_i, follows some distribution with probability density function $f_i(\cdot)$ and virtual valuation function $\phi_i(\cdot)$.
- All n bidders have values per unit in $[1, H]$.

Just like in the case of unit-demand bidders, the mechanism class we propose is once again inspired by t-level auctions introduced in [23, Corollary 5.6]. However, in this case the clinching function is not trivial and the arguments that establish an upper bound on the representation and error bounds of the class are quite more involved.

Definition 5.4 (Linear Deferred-Acceptance t-level Auctions ($\mathcal{LDA}$))**.** *Each bidder i faces t thresholds: $0 \le l_{i,0} \le l_{i,1} \le \cdots \le l_{i,t-1}$. This set of $t \cdot n$ numbers defines a Linear DA t-level auction with the following allocation and payment rules.*
Consider a valuation profile $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$:

1. *For each bidder i, let $t_i(v_i)$ denote the index of the largest threshold $l_{i,\tau}$ that lower bounds v_i (or -1, if $v_i < l_{i,0}$). We call $t_i(v_i)$ the level of bidder i.*
2. *Sort the bidders from lowest level to highest and, within a level, use a fixed lexicographical tie-breaking ordering $\prec$ to pick a winner.*
3. ***DA Allocation:** Start finalizing the bidders, from lowest level to highest. Every bidder is allocated, at the time that she is finalized, as many units as possible, provided that the units left afterwards are enough to fully satisfy the demands of the still active bidders.*
4. ***Payment Rule:** The unique one that renders truthful bidding a dominant strategy and charges zero to players that win no units: Every bidder, for every unit that she is allocated, pays for it the minimum she could have bid and still won that unit.*

Example 5.2. *There are 3 bidders, a, b and c. The thresholds are the same for every bidder: $l_{a,\cdot} = l_{b,\cdot} = l_{c,\cdot} = [2, 4, 6, 8, 10]$. Their values per unit are $\boldsymbol{v} = (v_1, v_2, v_3) = (4.5, 6.5, 9)$. Every bidder is interested in acquiring up to 5 units and there are 8 units available for sale. Finally, let $a \prec b \prec c$. Then:*

1. *Bidder a is finalized first and clinches no units because in order to satisfy the demands of all the still active bidders 10 units are required.*
2. *Bidder b is finalized second. She clinches 3 units, because the 5 units remaining afterwards are enough to fully satisfy the only remaining bidder. For every one of those units she pays a price of 4, because that is the minimum bid with which she would be at the same level as bidder a and then bidder b would win those units because of the tie-breaking rule $\prec$. So, her total payments are $3 \cdot 4 = 12$.*
3. *Bidder c is finalized last and clinches all 5 units that she wanted. For the first 3 of those units she pays a price of 4 per unit, because the reason that she won those units was that she was finalized after bidder a. For the last 2 of her units she pays a price of 6 per unit because the reason that she won those extra units was that she was finalized after bidder b. Thus, her total payments are $3 \cdot 4 + 2 \cdot 6 = 24$.*

Proposition 5.3. *Any linear DA auction, as described in definition 5.4, is a valid generalized DA auction.*

Proof. Let $0 \leq l_{i,0} \leq l_{i,1} \leq \cdots \leq l_{i,t-1}$ be the thresholds that bidder i faces in the unit-demand DA t-level auction. Formally, t the clinching and scoring functions that correspond to definition 5.4 are:

- scoring functions: $\sigma_i^{A_t}(b_i, \boldsymbol{b}_{N \setminus A_t}) = t_i(v_i)$
- clinching functions: $g_i^{A_t}(b_{N \setminus A_t}) = \min\left\{d_i, m - \min\left\{m, \sum_{j \in A_t \setminus \{i\}} d_j\right\}\right\}$

Just like in the case of unit-demand bidders, the scoring function of every bidder is weakly increasing in her bid since the thresholds she faces are weakly increasing and her level is simply the index of the highest level that is still lower than her bid. Also, the scoring function of every bidder does not depend on the bids of the other still active bidders.
The clinching function of every bidder does not depend on the bids of the other still active bidders. Also, it is easy to see that as the set of active bidders decreases, the amout of units that the still-active bidders clinch weakly increases: $\sum_{j \in A_t \setminus \{i\}} d_j$ strictly decreases as the set of active bidders decreases, therefore $\min\left\{m, \sum_{j \in A_t \setminus \{i\}} d_j\right\}$ weakly decreases.

Both the scoring and clinching functions adhere to all the constraints of definition 4.11, therefore any such auction is a valid generalized DA auction.

□

Now that we have established that the $\mathcal{LDA}$ mechanism class only contains valid generalized DA auctions, the next step is to determine the sample complexity of learning an auction from this class that has expected revenue that is a $(1-\epsilon)$ fraction of the optimal. To do this, we will bound separately the representation and *generalization* error of the class. The main idea is the same as in [23]: Find the value of t (number of thresholds) that balances optimally between a high representation and a high generalization error.

5.1.4 Generalization Error Bound

Theorem 5.2. *The pseudo-dimension of Linear DA t-level auctions with n bidders is $O(nt\log(nt))$.*

The idea is similar to [23], but some non-trivial modifications are required. First, we will introduce 2 standard results from learning theory:

Lemma 5.1 (Sauer's Lemma)**.** *Let $\mathcal{C}$ be a set of functions from $\mathcal{Q}$ to {0,1} with VC dimension d, and $\mathcal{S} \subseteq \mathcal{Q}$. Then:*

$$|\{\mathcal{S} \cap \{x \in \mathcal{Q} : c(x) = 1\} : c \in \mathcal{C}\}| \leq |S|^d \tag{5.1}$$

Lemma 5.2. *The set of linear separators in $\mathbb{R}^d$ has VC dimension $d+1$.*

Proof. Consider a set of samples S of size N which can be shattered by linear DA t-level auctions with revenue targets $(r^1, r^2, \ldots, r^N)$. We upper-bound the number of labelings of S possible using linear DA t-level auctions, which yields an upper bound on N.
For a fixed sample set of size N, we partition the auctions into *equivalence classes*, identically to the proof of [23, Theorem 3.3]. Across all auctions in an equivalence class, all comparisons between two thresholds or a threshold and a bid are resolved identically. Just like in [23, Theorem 3.3], for n bidders, t thresholds and a sample set of size N the number of equivalence classes are at most $(nN+nt)^{2nt}$. We now upper bound the number of distinct labelings any fixed equivalence class $\mathcal{C}$ of auctions can generate.
Consider a class $\mathcal{C}$ of equivalent auctions. The allocation and payment rules are relatively simple: The number of units a bidder i wins depends only on 3 things: the ordering of the bidders (by level), the fixed tie-breaking rule $\prec$ and the demands of all bidders and thus, for fixed demands, is a function only of bids and thresholds. This implies that, for every sample in S, all auctions in the class $\mathcal{C}$ result in the exact same allocation (number of units allocated to each bidder). This, along with Myerson payments, also imply that the payment of each winning bidder is a fixed sum of thresholds and coefficients that correspond to how many additional items he clinched because of that threshold:

$$p_{\mathcal{C},i}(v) = \sum_{j \leq t(v_i)} l_{i,j} \cdot h_{\mathcal{C},i}(\boldsymbol{v}, j) \tag{5.2}$$

where $h_{\mathcal{C},i}(v,j)$: How many additional units bidder i clinches in all auctions in $\mathcal{C}$, for the valuation profile v, because he is at least at level j. For any bidder i, the function $h_{\mathcal{C},i}()$ is the same across all samples in set S, for any auction in $\mathcal{C}$.
Now, we encode each $\mathcal{A} \in \mathcal{C}$ and sample $\boldsymbol{v}^j$ as an $(nt+1)$-dimensional vector as follows: Let $x^{\mathcal{A}}_{i,\tau}$ encode the value of $l_{i,\tau}$ in the auction $\mathcal{A}$. Define $x^{\mathcal{A}}_{nt+1} = 1$ for every $\mathcal{A} \in \mathcal{C}$. Define

$y^j_{i,\tau} = h_{\mathcal{C},i}(v,j)$ (for how many units bidder i is paying her τ-th threshold). Finally, define $y^j_{nt+1} = -r^j$. The point is that, for every auction $\mathcal{A}$ in the class $\mathcal{C}$ and sample $\boldsymbol{v}^j$,

$$x^{\mathcal{A}} \cdot y^j \geq 0 \tag{5.3}$$

if and only if $\text{Rev}(\mathcal{A}, \boldsymbol{v}^j) \geq r^j$. Thus, the number of distinct labelings of the samples generated by auctions in $\mathcal{C}$ is bounded above by the number of distinct sign patterns on N points in $\mathbb{R}^{nt+1}$ generated by all linear separators (The y^j-vectors are constant across $\mathcal{C}$ and can be viewed as m fixed points in $\mathbb{R}^{nt+1}$; each auction $\mathcal{A} \in \mathcal{C}$ corresponds to the vector $x^{\mathcal{A}}$ of coefficients.). Applying Sauer's Lemma and Lemma 5.2, LDA t-level auctions can generate at most N^{nt+2} labelings per equivalence class, and hence at most $(nN + nt)^{3nt+2}$ distinct labelings in total. This imposes the restriction:

$$2^N \leq (nN + nt)^{3nt+2} \tag{5.4}$$

Solving for N yields the desired bound. □

5.1.5 Representation Error Bound

Theorem 5.3. *Consider the environment described above (budget-additive bidders). Suppose F is a product distribution with support $[1, H]^n$. Then, for $t = \lceil \frac{2}{\epsilon} \rceil + \log_{1+\frac{\epsilon}{2}} H = \Theta\left(\frac{1}{\epsilon} + \log_{1+\epsilon} H\right)$, the class of linear DA t-level auctions ($\mathcal{LDA}$) contains an auction with expected revenue at least $(1-\epsilon)$ times the optimal one.*

The proof is in the same vein as the proof of [23, Theorem 5.4], but once again some non-trivial modifications are required.

Proof. Consider a fixed bidder i. We define t thresholds for i, bucketing her by her virtual value, and prove that the t-level DA Auction $\mathcal{A}$ induced by these thresholds for each bidder closely approximates the expected revenue of the optimal auction, Myerson, denoted $\mathcal{M}$.

- Set $l_{i,0} = \phi^{-1}(0)$, bidder i's monopoly reserve.
- For $\tau \in [1, \lceil 2/\epsilon \rceil]$, let $l_{i,\tau} = \phi_i^{-1}(\tau \cdot \frac{\epsilon}{2})$ $\quad(\phi_i \in [0,1])$
- For $\tau \in [\lceil 2/\epsilon \rceil, \lceil 2/\epsilon \rceil + \lceil \log_{1+\frac{\epsilon}{2}} H \rceil]$, let $l_{i,\tau} = \phi_i^{-1}(d(1+\frac{\epsilon}{2})^{\tau - \lceil \frac{2}{\epsilon} \rceil})$ $\quad(\phi_i > 1)$

Let $\mathcal{A}$ denote the corresponding LDA t-level auction. Fix an arbitrary valuation profile $\boldsymbol{v}$. We will compare the Virtual Social Welfare of $\mathcal{A}$ to that of $\mathcal{M}$. For every unit allocated to some bidder i, up to her demand d_i, the contribution of that unit to the virtual social welfare is $\phi_i(v_i)$ (i.e. equal to that bidder's virtual value).

- By its definition, $\mathcal{M}$ is optimal with respect to the virtual social welfare and allocates units that correspond to the m highest *duplicated* virtual values, where each virtual value $\phi_i(v_i)$ has been duplicated d_i times (the demand of the corresponding bidder), provided those m duplicated virtual values are positive.
- In a similar vein, the allocation of $\mathcal{A}$ is lexicographically optimal with respect to the levels, rather than the exact virtual values. Since the allocation of $\mathcal{M}$ is the actual optimal, it is also optimal with respect to the levels. Thus, the level of the unit with the j-th highest virtual value in the allocation of $\mathcal{A}$ is the same as the unit with the j-th highest virtual value in the allocation of $\mathcal{M}$ (otherwise $\mathcal{A}$ wouldn't be optimal with respect to the levels).

- The 2 mechanisms always allocate the same number of units: the minimum between m and the number of duplicated virtual values ≥ 0.
- Then, by an accounting argument identical to [23, Theorem 3.5] summing up over all units completes the proof:
  - Consider a fixed valuation profile $\boldsymbol{v}$.
  - Let i^* and i' denote the i-th highest unit in the allocation of $\mathcal{A}$ and $\mathcal{M}$ respectively.
  - Both auctions only allocate to non-negative (ironed) virtual values.
  - Let τ be the level of i^*
  - If there is no tie at that level, then both Mechanisms allocate that same unit.
  - When there is a tie at level τ, the virtual value of the unit allocated by $\mathcal{A}$ is close to that of $\mathcal{M}$:
    * If $\tau \in [0, \lceil 2/\epsilon \rceil]$, then $\phi_{i'}(v_{i'}) - \phi_{i^*}(v_{i^*}) \leq \frac{\epsilon}{2}$.
    * If $\tau \in [\lceil 2/\epsilon \rceil, \lceil 2/\epsilon \rceil + \lceil \log_{1+\frac{\epsilon}{2}} H \rceil]$, then $\frac{\phi_{i^*}(v_{i^*})}{\phi_{i'}(v_{i'})} \geq 1 - \frac{\epsilon}{2}$.
- These facts imply that:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{v}}[\mathrm{Rev}_i(\mathcal{A}, \boldsymbol{v})] = \mathbb{E}_{\boldsymbol{v}}[\phi_{i^*}(v_{i^*})] &\geq \left[\left(1 - \frac{\epsilon}{2}\right) \mathbb{E}_{\mathrm{v}}[\phi_{i'}(v_{i'})] - \frac{\epsilon}{2}\right] \\ &\geq \left(1 - \frac{\epsilon}{2}\right) \cdot \mathbb{E}_{\mathrm{v}}[\phi_{i'}(v_{i'})] - \frac{\epsilon}{2} = \left(1 - \frac{\epsilon}{2}\right) \mathbb{E}_{\mathrm{v}}[Rev_i(\mathcal{M})] - \frac{\epsilon}{2} \end{aligned} \tag{5.5}$$

  where $\mathrm{Rev}_i(\mathcal{A}, \boldsymbol{v})$: The contribution to the expected social welfare of $\mathcal{A}$ of the unit with the i-th highest level, breaking ties according to $\prec$.
- Finally, because all bidders' valuations are bounded in the range $[1, H]$, at price 1 any unit would be sold. Thus, $\frac{\epsilon}{2}$ is at most a $\frac{\epsilon}{2}$-fraction of the expected revenue per unit at the optimal reserve price.
- Summing over the m units with the highest levels and virtual values for $\mathcal{A}$ and $\mathcal{M}$ and using linearity of expectations completes the proof.

□

Combining the previous 2 theorems (5.2 and 5.3) with theorem 5.1 and observing that the maximum possible revenue of any auction for valuations in $[1, H]$ and m units is mH immediately yields the following corollary:

Corollary 5.2. *With probability $1 - \delta$, the empirical revenue maximizer for a sample of size N of the class of Linear DA t-level auctions is a $1 - O(\epsilon)$-approximation to Myerson for m units and n bidders whose valuations are in $[1, H]$, for $t = O(\frac{1}{\epsilon} + \log_{1+\epsilon} H)$ and*

$$N = O\left(\left(\frac{Hm}{\epsilon}\right)^2 \left(nt \log(nt) \ln \frac{Hm}{\epsilon} + \ln \frac{1}{\delta}\right)\right) = \widetilde{O}\left(\frac{H^2 m^2 n}{\epsilon^3}\right) \tag{5.6}$$

Intuitively, this result makes sense: For bidders with linear valuations functions, it was proven in [18] that VCG is implementable as a DA auction. Therefore Myerson, if viewed as the virtual surplus maximization mechanism, is also implementable as a DA auction. Thus, the problem of learning the revenue-maximizing DA auction for this setting reduces to learning the bidders' virtual valuation functions and then implementing the resulting mechanism for those inferred virtual valuation functions as a DA auction.

5.2 Extending the Results to Polymatroid Constraints

A natural generalization of multi-unit auctions with identical goods are the settings where the set of feasible outcomes is defined by a polymatroid constraint. Pictorially, in any such setting the space of feasible solutions constitutes a *polytope*. Formally:

Definition 5.5 (Polymatroid Environment)**.** *For n agents in a polytope environment, the set of feasible environments is defined via a given submodular function $h : 2^n \to \mathbb{R}_+$, as follows:*

$$P_h = \left\{ l \in \mathbb{N}^n | \sum_{i \in S} l_i \leq h(S) \; \forall S \subseteq N \right\} \tag{5.7}$$

As a first example, a multi-unit auction with m available units can also be viewed as a polymatroid environment: The submodular constraint function in that case is the constant function $h(S) = m$ for every $S \subseteq N$, which can be naturally interpreted as "no subset of players can be allocated more units than the total number of units available".
For a more interesting example consider a *keyword sponsored search auction*, where the agents are competing for a sequence of $m < n$ advertising slots, and each slot j has a click-through rate r_j. Then, if bidder i's value per click is v_i, her total value for slot j is $r_j \cdot v_j$. It has been experimentally proven that higher slots have higher click-through rates, i.e. $r_1 \geq r_2 \cdots \geq r_m$. Now, the constraint can be naturally interpreted as "no group of agents can be allocated a click-through rate higher than the total click through rate, and no group of $k \leq m$ agents can be allocated more than the total click-through rate of the k highest slots". Formally, the polymatroid constraint in this setting is defined by the submodular function $h(S) = \sum_{j-1}^{|S|} r_j$. For more motivating examples behind polymatroid constraints we refer the interested reader to [24].
We will shift our attention to perhaps the most natural extension of the multi-unit auctions we studied previously. Once again every bidder's value "per level of service" is bounded in some range $[1, H]$, but now the set of feasible outcomes is defined by some polymatroid constraint. Formally:

The Setting (Polymatroid Constraints, Additive Valuations)

- n bidders
- every bidder is additive, up to her demand:
  For every bidder i, her value for acquiring a "level of service" x_i is $x_i \cdot v_i$, where v_i: Her value per level of service.
- Every bidder i's value per unit, v_i, follows some distribution with probability density function $f_i(\cdot)$ and virtual valuation function $\phi_i(\cdot)$
- All n bidders have values per unit on $[1, H]$.
- The set of feasible allocations is defined via a given submodular function
  $h : 2^n \to \mathbb{R}_+$, as follows:

$$P_h = \left\{ x \in \mathbb{R}^n | \sum_{i \in S} x_i \leq h(S) \; \forall S \subseteq N \right\}$$

For this environment, the first mechanism class we propose are the polymatroid DA t-level auctions. As the name suggests, this mechanism class is a combination of polymatroid DA auctions introduced in [18] and t-level auctions introduced in [23].

Definition 5.6 (Polymatroid DA t-level auction ($\mathcal{PDA}$)). *In a polymatroid DA t-level auction every bidder i faces t thresholds, $0 \leq l_{i,0} \leq l_{i,1} \leq \cdots \leq l_{i,t-1}$. Let $t_i(b_i)$ denote the index of the largest threshold $l_{i,\tau}$ that lower bounds b_i (or -1, if $b_i < l_{i,0}$). This set of $t \cdot n$ numbers defines a Polymatroid DA t-level auction with the following sets of scoring and clinching functions: Consider a bid profile $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$:*

- *Scoring function:* $\sigma_i^{A_t}(b_i, b_{N \setminus A_t}) = t_i(b_i)$
- *Clinching function:* $g_i^{A_t}(b_{N \setminus A_t}) = \begin{cases} h(A_t) - h(A_t \setminus \{i\}), \text{if } t_i(b_i) \geq 0 \\ 0, else \end{cases}$

This DA auction has a very simple description: At every stage the bidder that is finalized is the one with the lowest level, among the ones that are still active (breaking ties according to the tie-breaking rule $\prec$). The clinching function is more interesting: At each stage t each bidder i has clinched a level of service equal to her marginal contribution to the value of $h(A_t)$. Because h is submodular, the marginal contribution weakly increases as A_t shrinks, so this is valid clinching function. Finally, it is easy to see that the scoring function of every bidder does not depend on the bids of the other still-active players and is weakly increasing in her bid, so this is valid scoring function. Thus, the $\mathcal{PDA}$ class only contains valid generalized DA auctions.

Proposition 5.4. *Any polymatroid DA t-level auction, as described in definition 5.6, is a valid generalized DA auction.*

Our approach is similar to the one for the multi-unit auction: We will establish representation and generalization error bounds, and then combine them using Pollard's uniform convergence theorem.

Theorem 5.4 (Representation Error Bound). *Consider the environment described above (linear bidders, polymatroid constraints). Suppose F is a product distribution with support in $[1, H]^n$. Provided $t = O\left(\frac{H}{\epsilon}\right)$, there exists a polymatroid DA t-level auction with expected revenue at least a $1 - \epsilon$ fraction of the optimal expected revenue.*

Proof. Consider a fixed bidder i. We define t thresholds for i bucketing her by her virtual value, and prove that the t-level DA Auction $\mathcal{A}$ induced by these thresholds for each bidder closely approximates the expected revenue of the optimal auction $\mathcal{M}$.

- Set $l_{i,0} = \phi^{-1}(0)$, bidder i's monopoly reserve.
- For $\tau \in [1, \lceil 2H/\epsilon \rceil]$, let $l_{i,\tau} = \phi_i^{-1}(\tau \cdot \epsilon/2)$

We will compare the expected virtual social welfare of the polymatroid DA t-level auction with those thresholds against the expected social welfare of Myerson mechanism. Consider a fixed valuation profile $\boldsymbol{v}$. Let $\phi(\boldsymbol{v}) = (\phi_1(v_1), \phi_2(v_2), \cdots, \phi_n(v_n))$ be the vector of all virtual values in that valuation profile. Myerson is essentially solving the following linear problem:

$$\begin{aligned} &\max \phi(\boldsymbol{v}) \cdot x \\ &x \in P_h \end{aligned} \tag{5.8}$$

On the other hand, our DA auction leads on the exact same allocation as Edmonds' greedy algorithm (already proven in [5]), for the vector $t(\boldsymbol{v}) = (t_1(v_1), t_2(v_2), \cdots, t_n(v_n))$, i.e. it solves:

$$\begin{aligned} \max t(\boldsymbol{v}) \cdot x \\ x \in P_h \end{aligned} \tag{5.9}$$

But for any bidder i, at any point where her virtual value increases by exactly $\epsilon/2$ (past 0), her level increases by 1. Thus, it holds that:

- If $t_i(\mathrm{v}_i) < 0$, then $\phi_i(\mathrm{v}_i) < 0$.
- If $t_i(\mathrm{v}_i) \geq 0$, then $\phi_i(\mathrm{v}_i) \in \left[t_i(\mathrm{v}_i) \cdot \frac{\epsilon}{2}, t_i(\mathrm{v}_i) \cdot \frac{\epsilon}{2} + \frac{\epsilon}{2}\right)$

So, since the solution of our auction is optimal with respect to the vector $t(\boldsymbol{v}) = (t_1(v_1), t_2(v_2), \cdots, t_n(v_n))$, by the definition of the levels it is also optimal with respect to the vector

$$\begin{aligned} \widetilde{\phi}(\boldsymbol{v}) &= \frac{\epsilon}{2}(t_1(v_1), t_2(v_2), \cdots, t_n(v_n)) \\ &= \Big(\frac{\epsilon}{2} \cdot t_1(v_1), \frac{\epsilon}{2} \cdot t_2(v_2), \cdots, \frac{\epsilon}{2} \cdot t_n(v_n)\Big) \\ &= (\phi_1(v_1) - \epsilon_1, \phi_2(v_2) - \epsilon_2, \cdots, \phi_n(v_n) - \epsilon_n), \end{aligned} \tag{5.10}$$

where: $\epsilon_i \in [0, \epsilon/2)$ for all $i \in N$ such that $\phi_i(v_i) \geq 0$.
The loss in approximation for those players that their virtual value (and level) is less than zero doesn't matter, because for the allocation of both our algorithm and the optimal one it holds that $x_i = 0$ for any player i such that her virtual value (and therefore level) is below 0.
The reason why the allocation of our algorithm is also optimal with respect to the vector $\widetilde{\phi}(\boldsymbol{v})$ is because we multiplied all the coordinates of vector $t()$v with the same positive number, so for any allocation x it holds that $x \cdot \widetilde{\phi}(\boldsymbol{v}) = \frac{\epsilon}{2} x \cdot t(\boldsymbol{v})$.
Let x^*, x be the allocations of Myerson and our mechanism respectively. For the expected virtual social welfare of our mechanism we have:

$$\begin{aligned} \mathbb{E}[\phi(\boldsymbol{v}) \cdot x] &\geq \mathbb{E}[\widetilde{\phi}(\boldsymbol{v}) \cdot x] \\ &\geq \mathbb{E}[\widetilde{\phi}(\boldsymbol{v}) \cdot x^*] \\ &= \mathbb{E}[\sum_i (\phi_i(v_i) - \epsilon_i) \cdot x_i^*] \\ &= \mathbb{E}[\sum_{i: x_i^* > 0} (\phi_i(v_i) - \epsilon_i) \cdot x_i^*] \\ &\geq \mathbb{E}[\sum_{i: x_i^* > 0} \Big(\phi_i(v_i) - \frac{\epsilon}{2}\Big) \cdot x_i^*] \\ &\geq \mathbb{E}[\sum_{i: x_i^* > 0} \phi_i(v_i) \Big(1 - \frac{\epsilon}{2}\Big) \cdot x_i^*] \\ &= \Big(1 - \frac{\epsilon}{2}\Big) \mathbb{E}[\sum_{i: x_i^* > 0} \phi_i(v_i) \cdot x_i^*] \\ &= \Big(1 - \frac{\epsilon}{2}\Big) \mathbb{E}[\sum_i \phi_i(v_i) \cdot x_i^*] \end{aligned} \tag{5.11}$$

Where the first inequality follows from the fact that $\phi(\boldsymbol{v}) \geq \widetilde{\phi}(\boldsymbol{v})$, the second inequality follows from the fact that x is optimal with respect to the vector $\widetilde{\phi}(\boldsymbol{v})$, the first equality follows from

the definition of $\widetilde{\phi}(\boldsymbol{v})$ and the second one follows from the fact that $x_i^* \geq 0 \forall i$. The third inequality follows from the fact that as discussed earlier, for any player with non-negative virtual value, $\epsilon_i \in [0, \epsilon/2)$, the forth inequality follows from the fact that at price 1 everyone would be interested in buying as high a level of service as possible, thus $\epsilon/2$ is always at most a $\epsilon/2$-fraction of a winning bidder's expected payment per unit. Finally, the forth equality follows from linearity of expectations and the last one follows once again from the fact that $x_i^* \geq 0 \; \forall i$.
Combining equation 5.11 with the fact that for every mechanism its expected revenue is equal to its expected virtual social welfare, it follows immediately that the expected revenue of our polymatroid DA t-level auction is at least a $1 - \frac{\epsilon}{2}$ of the expected revenue of Myerson. □

Now, it suffices to establish a generalization error bound. Once again this will be accomplished by bounding the pseudo-dimension of the class.

Theorem 5.5. *The pseudo-dimension of polymatroid DA t-level auctions with n bidders is $O(nt \log(nt))$.*

The proof is identical to that of theorem 5.2. Again, for a fixed sample set and number of levels t we can break the auctions into equivalence classes in the same way. Then, for any sample, all auctions in an equivalence class will result in the same allocation and payment identity. Thus, for any sample we can encode the common outcome of all auctions in an equivalence class on that sample as an $(nt+1)$-dimensional array and then repeat the same argument upper bounding the number of possible labelings on that sample by the number of possible labelings for that sample of linear separators in the $(nt+1)$-dimensional space.
Finally, it suffices to observe that for a fixed polymatroid constraint defined by some submodular function h on a set $\mathcal{N}$ bidders the maximum possible revenue of any mechanism is $H \cdot h(\mathcal{N})$, as this is the maximum possible social welfare, for any auction. This observation, along with theorems 5.4, 5.5 and 5.1 immediately yield:

Corollary 5.3. *With probability at least $1-\delta$, for any environment with polymatroid constraints defined by some submodular funciton h, the empirical revenue maximizer for a sample set of size N of the class of PDA t-level auctions is a $1 - O(\epsilon)$-fraction to Myerson for n linear bidders whose valuations per unit are in $[1, H]$, for $t = O\left(\frac{H}{\epsilon}\right)$ and*

$$N = O\left(\left(\frac{H \cdot h(S)}{\epsilon}\right)^2 \left(nt \log(nt) \ln \frac{H \cdot h(S)}{\epsilon} + \ln \frac{1}{\delta}\right)\right) = \widetilde{O}\left(\frac{H^3 h(S)^2 n}{\epsilon^3}\right),$$

where $H \cdot h(S)$ is the maximum social welfare that could be possible under those polymatroid constraints, and therefore is an upper bound on the revenue achievable by any mechanism in the class of PDA t-level auctions, for the given distribution F.

5.2.1 Improving the Previous Result

Even though the extension of DA auctions to settings with polymatroid constraints of the previous section was interesting, the approach of that subsection was short sighted: Intuitively, in the previous section the main idea was to learn, for every bidder, a discretization of her virtual valuation function that is ϵ-accurate *additively* over the whole range of positive virtual values. But in order to achieve expected revenue at least $(1-\epsilon)$ times the optimal one, it is only necessary to learn a discretization of all the bidders' virtual valuation functions that is ϵ-accurate *multiplicatively* over that range. As a result, in the previous section the complexity

of the t-levels required per bidder, and therefore samples, was blown up by about a factor of H.
However, if we simply attempt to reduce the number of levels per bidder, this introduces a new challenge: Unlike multi-unit auctions, for polymatroid constraints the fact that an auction is optimal with respect to the levels does not imply that it is also optimal with respect to the actual virtual valuations of the bidders, because the two quantities are no longer linearly correlated (plus some small perturbation). To tackle this issue, we now supplied each level of a DA t-level auction with one additional number, which we named the *level score*. Level scores are learned on the samples, just like the level thresholds. The mechanism now does not choose an allocation that is maximal with respect to the indexes of the levels, but with respect to the level scores. Thus, if the level scores for any level are ϵ-close multiplicatively to the virtual value at that point, the virtual social welfare of the final allocation will also be ϵ-close to the virtual social welfare of the optimal allocation.
We call the mechanism class we just described extended polymatroid DA t-level auctions. Formally:

Definition 5.7 (Extended Polymatroid DA t-level auction ($\mathcal{EPDA}$))**.** *In an extended polymatroid DA t-level auction every bidder i faces t thresholds,* $0 \leq l_{i,0} \leq l_{i,1} \leq \cdots \leq l_{i,t-1}$*. Let* $t_i(b_i) = \tau_i$ *denote the index of the largest threshold* $l_{i,\tau}$ *that lower bounds* b_i *(or* -1*, if* $b_i < l_{i,0}$*). For every one of the t levels of every bidder, there is an associated level score satisfying:* $r_{i,-1} = -1$, $0 \leq r_{i,0} \leq r_{i,1} \leq \cdots \leq r_{i,t}$*. When a bidder is at her level* τ_i*, her corresponding level score is* r_{i,τ_i}*. This set of* $2t \cdot n$ *parameters defines an extended polymatroid DA t-level auction with the following sets of scoring and clinching functions: Consider a bid profile* $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$*:*

- *Scoring function:* $\sigma_i^{A_t}(b_i, b_{N \setminus A_t}) = r_{i,t_i(b_i)} = r_{i,\tau_i}$
- *Clinching function:* $g_i^{A_t}(b_{N \setminus A_t}) = \begin{cases} h(A_t) - h(A_t \setminus \{i\}), \textit{if } t_i(b_i) \geq 0 \\ 0, \textit{else} \end{cases}$

The algorithmic description of this mechanism class is very similar to the one of polymatroid DA t-level auctions, introduced in the previous subsection: At every stage the bidder that is finalized is the one with the lowest level score, among the ones that are still active (breaking ties according to the tie-breaking rule $\prec$). At each stage t each bidder i has clinched a level of service equal to her marginal contribution to the value of $h(A_t)$. Because h is submodular, the marginal contribution weakly increases as A_t shrinks, so this is valid clinching function. The scoring function of every bidder does not depend on the bids of the other still active bidders, and is weakly increasing in her bid: As a bidder increases her bid, her level weakly increases, which in turn implies that her level score weakly increases, so this is once again a valid scoring function. Thus, the $\mathcal{EPDA}$ class only contains valid generalized DA auctions.

Proposition 5.5. *Any extended polymatroid DA t-level auction, as described in definition 5.7, is a valid generalized DA auction.*

Once again, in order to establish a sample complexity bound for this mechanism class our approach will be to separately bound its representation and generalization errors. However, now our definition has diverged so far from the original definition of t-level auctions that there is no apparent way of generalizing the pseudo-dimension arguments of [23]. In order to bound the pseudo-dimension of $\mathcal{EPDA}$ we will now resort to the sample complexity framework of Balcan ([12]).

Theorem 5.6. *The class of extended polymatroid DA t-level auctions is $(2nt, n^2t^2)$-delineable.*

Proof. For the parameter space $\mathcal{P}$ we have:

- $(l_{i,0}, l_{i,1}, \ldots, l_{i,t-1}) \in \mathbb{R}_+^t \ \forall i \in [n]$
- $(r_{i,0}, r_{i,1}, \ldots, r_{i,t-1}) \in \mathbb{R}_+^t \ \forall i \in [n]$
- $(l_{i,0} \leq l_{i,1} \leq \cdots \leq l_{i,t-1}) \in \mathbb{R}_+^t \ \forall i \in [n]$
- $(r_{i,0} \leq r_{i,1} \leq \cdots \leq r_{i,t-1}) \in \mathbb{R}_+^t \ \forall i \in [n]$

Fix a valuation profile $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$:

- There are in total $t \cdot n$ level scores.
- Once we have have determined whether we are on the positive or negative side of all $(nt)(n(t-1))/2 = O((nt)^2)$ hyperplanes of the form $r_{i,t_1} \geq r_{j,t_2}$ for all $i \neq j$ and all $t_1, t_2 \in [t]$, we have determined a unique total ordering on all nt level scores (breaking ties in level scores by some fixed lexicographical ordering on the bidders, $\prec$).
- Thus, there are at most $(nt)^2$ hyperplanes splitting the parameter space in connected components, in each of whom the total ordering of all nt level scores remains constant.
- To determine the level of a single bidder i, we need to check whether we are on the positive or negative side of all t hyperplanes of the form: $l_{i,\tau} \leq v_i$, $\tau \in [t]$.The answer is the unique index τ_i for which it holds $l_{i,\tau_i} \leq v_i$ and $l_{i,\tau_i+1} > v_i$.
- Multiplying by n for all bidders, there are at most nt hyperplanes splitting the parameter space into connected components, in each of whom the levels of all bidders, for a fixed valuation profile $\boldsymbol{v}$, are fixed.
- Thus, in each connected component of the parameter space minus both the hyperplanes for the total ordering of all nt "level scores" and the hyperplanes which determine the level of each bidder, it holds:
  1. The total ordering of all levels scores $r_{i,j}$, $i \in [n], j \in [t]$ is fixed.
  2. The level of each bidder is fixed.
  3. The total ordering of all bidders by their level scores is fixed.
- We will restrict our attention to one such connected component.
- Now it suffices to observe that once all the above orderings are fixed, for a fixed valuation profile $\boldsymbol{v}$ every mechanism in that connected component results in the same allocation:
  1. Every mechanism starts finalizing the bidders one-by-one, from lowest level score to highest.
  2. For every mechanism in one such a connected component, the total ordering of all bidders is the same, so bidders are finalized in the exact same order.
  3. The allocation of every bidder depends only on the set of active bidders when she is finalized.
  4. Thus, because in every mechanism in that connected component, for a fixed valuation profile the ordering in which bidders are finalized is the same, the set of active bidders when every one is finalized is also the same and therefore the resulting allocation of every bidder is the same.

- Finally, for a fixed valuation profile in that connected component, the payment of every bidder depends linearly on her thresholds, and does not depend (as long as we remain in that connected component) on the level scores:

  Let τ_i be the actual level of bidder i. Let $x_i(t_i)$ be the allocation of bidder i, if her level was t_i, and let $x_i(0) = 0$. Holding the ordering of all nt level scores fixed (they are in each connected component), and holding the bids of all other players except i fixed to b_{-i}, as i increases her bid, her level increases, which increases her level score, which in turn increases her relative position on the ordering of bidders by level scores, which finally increases her allocation. Then, Myerson payments dictate that the actual payments of bidder i are:

$$p_i(\boldsymbol{v}) = \sum_{j=1}^{\tau_i} \underbrace{(x_i(j) - x_i(j-1))}_{\text{Allocation because of level j}} \cdot \underbrace{l_{i,j}}_{\text{Min. bid for level j}} \tag{5.12}$$

- In each such connected component the levels of all the bidders and the ordering of all nt level scores is fixed, so both all the τ_i's and the differences $x_i(j) - x_i(j-1)$ are fixed.
- This along with equation 5.12 and the fact that the total revenue is simply the sum of the payments of all bidders imply that for a fixed valuation profile, the revenue of the mechanism in each such connected component is linear in the parameters of the mechanism (as long as we are in a connected component, the revenue depends linearly on the thresholds $l_{i,j}$, and does not depend on the level scores at all).
- Thus, for a fixed valuation profile $\boldsymbol{v}$, the total number of hyperplanes splitting the parameter space into regions where the profit from every bidder is linear is:

$$\underbrace{O((nt)^2)}_{\text{For the scores}} + \underbrace{O(nt)}_{\text{For the levels}} = O((nt)^2) \tag{5.13}$$

□

Combining the above theorem with Balcan's theorem 4.2 immediately yields:

Theorem 5.7. *The pseudo-dimension of the class of extended polymatroid DA t-level auctions is $O(nt\log(nt))$.*

This immediately results in a generalization error bound for the mechanism class. The only thing left now is to establish a representation error bound. The approach will be similar to the one we took for polymatroid auctions in theorem 5.5 but some additional arguments are required.

Theorem 5.8 (Representation Error Bound)**.** *Consider the environment described above (linear bidders, polymatroid constraints). Suppose F is a product distribution with support in $[1, H]^n$. Provided $t = O\left(\frac{1}{\epsilon} + \log_{1+\epsilon} H\right)$, there exists an extended polymatroid DA t-level auction with expected revenue at least $\left(1 - \frac{\epsilon}{2}\right)$ times the optimal expected revenue.*

Proof. Consider a fixed bidder i. We define t thresholds for i bucketing her by her virtual value and their corresponding thresholds, and prove that the t-level DA Auction $\mathcal{A}$ induced by these thresholds and scores for each bidder closely approximates the expected revenue of the optimal auction $\mathcal{M}$.

- Set $l_{i,0} = \phi^{-1}(0)$, bidder i's monopoly reserve and $r_{i,0} = 0$.
- For $\tau \in [1, \lceil 2/\epsilon \rceil]$, let $l_{i,\tau} = \phi_i^{-1}(\tau \cdot \epsilon/2)$ $\qquad (\phi_i \in (0,1])$
- For $\tau \in [1, \lceil 2/\epsilon \rceil]$, let $r_{i,\tau} = \tau$
- For $\tau \in [\lceil 2/\epsilon \rceil, \lceil 2/\epsilon \rceil + \lceil \log_{1+\frac{\epsilon}{2}} H \rceil]$, let $l_{i,\tau} = \phi_i^{-1}((1+\frac{\epsilon}{2})^{\tau - \lceil \frac{2}{\epsilon} \rceil})$ $\qquad (\phi_i > 1)$
- For $\tau \in [\lceil 2/\epsilon \rceil, \lceil 2/\epsilon \rceil + \lceil \log_{1+\frac{\epsilon}{2}} H \rceil]$, let $r_{i,\tau} = \lceil \frac{2}{\epsilon} \rceil \left(1 + \frac{\epsilon}{2}\right)^{\tau - \lceil \frac{2}{\epsilon} \rceil}$

For any bidder i, from the definition of the level scores it holds:

- If $\phi_i(v_i) < 0$ then $r_i = -1$.
- If $\phi_i(v_i) \in [0,1]$, then any time that bidder's virtual value increases by $\epsilon/2$, her level score increases by 1, so: $\phi_i(v_i) \in \left[\frac{\epsilon}{2} \cdot r_i(v_i), \frac{\epsilon}{2} \cdot r_i(v_i) + \frac{\epsilon}{2}\right)$
- Finally if $\phi_i(v_i) > 1$ then $\phi_i(v_i) \in \left[\frac{\epsilon}{2} \cdot r_i(v_i), \frac{\epsilon}{2} \cdot r_i(v_i) \left(1 + \frac{\epsilon}{2}\right)\right)$

The loss in approximation ratio from players with negative virtual value (and level score) doesn't matter, because for both our mechanism, and Myerson, their allocation will be 0. For any player i with $\phi_i(v_i) \geq 0$ it holds:

- If $\phi_i(v_i) \in [0,1]$ then $\frac{\epsilon}{2} \cdot r_i(v_i) \in \left[\phi_i(v_i) - \frac{\epsilon}{2}, \phi_i(v_i)\right]$
- If $\phi_i(v_i) \geq 1$ then $\frac{\epsilon}{2} \cdot r_i(\mathrm{v}_i) \in \left[\phi_i(\mathrm{v}_i) \cdot \left(1 - \frac{\epsilon}{2}\right), \phi_i(\mathrm{v}_i)\right]$

A final important remark is that for any bidder i with non-zero allocation by Myerson, her expected payment "per unit" is at least 1, because all bidder values are bounded in [1,H], and thus at price 1 everyone would be interested in purchasing as high an allocation as possible. So, for any such bidder i it holds $\phi_i(v_i) - \frac{\epsilon}{2} \geq \left(1 - \frac{\epsilon}{2}\right) \phi_i(v_i)$.
We will compare the expected virtual social welfare of the EPDA t-level auction with those thresholds and level scores against the expected virtual social welfare of Myerson. Consider a fixed valuation profile v. Let $\phi(\boldsymbol{v}) = (\phi_1(v_1), \phi_2(v_2), \cdots, \phi_n(v_n))$ be the vector of all virtual values in that valuation profile, and let $r(\boldsymbol{v}) = (r_1(v_1), r_2(v_2), \cdots, r_n(v_n))$ be the corresponding vector of all "level scores" in that valuation profile. Myerson is essentially solving the following linear problem:

$$\begin{aligned} &\max \phi(\boldsymbol{v}) \cdot x \\ &x \in P_h \end{aligned} \tag{5.14}$$

On the other hand, our DA auction leads on the exact same allocation as Edmonds' greedy algorithm for optimization over polymatroids (already proven in [18]), for the vector $r(\boldsymbol{v}) = (r_1(v_1), r_2(v_2), \cdots, r_n(v_n))$, i.e. it solves:

$$\begin{aligned} &\max r(\boldsymbol{v}) \cdot x \\ &x \in P_h \end{aligned} \tag{5.15}$$

Let x^*, x be the final allocations of Myerson and our EPDA t-level auction, respectively. x is optimal with respect to the vector $r(\boldsymbol{v})$, so it is also optimal with respect to the vector:

$$\begin{aligned} \frac{\epsilon}{2} r(\boldsymbol{v}) &= \left(\frac{\epsilon}{2} \cdot r_1(v_1), \frac{\epsilon}{2} \cdot r_2(v_2), \ldots, \frac{\epsilon}{2} \cdot r_n(v_n)\right) \\ &= (\phi_1(v_1) - \epsilon_1, \phi_2(v_2) - \epsilon_2, \ldots, \phi_n(v_n) - \epsilon_n) = \widetilde{\phi}(\boldsymbol{v}) \end{aligned} \tag{5.16}$$

where as we explained before, for any bidder i with non-zero allocation either on x or x^* it holds that $\epsilon_i \leq \frac{\epsilon}{2} \cdot \phi_i(v_i)$.
So, for the expected virtual social welfare of our mechanism, we have:

$$\begin{aligned}
\mathbb{E}[\phi(\boldsymbol{v}) \cdot x] &\geq \mathbb{E}[\widetilde{\phi}(\boldsymbol{v}) \cdot x] \\
&\geq \mathbb{E}[\widetilde{\phi}(\boldsymbol{v}) \cdot x^*] \\
&= \mathbb{E}[\sum_i (\phi_i(v_i) - \epsilon_i) \cdot x_i^*] \\
&= \mathbb{E}[\sum_{i: x_i^* > 0} (\phi_i(v_i) - \epsilon_i) \cdot x_i^*] \\
&\geq \mathbb{E}[\sum_{i: x_i^* > 0} \phi_i(v_i) \left(1 - \frac{\epsilon}{2}\right) \cdot x_i^*] \\
&= \left(1 - \frac{\epsilon}{2}\right) \mathbb{E}[\sum_{i: x_i^* > 0} \phi_i(v_i) \cdot x_i^*] \\
&= \left(1 - \frac{\epsilon}{2}\right) \mathbb{E}[\sum_{i: x_i^* > 0} \phi_i(v_i) \cdot x_i^*] + \left(1 - \frac{\epsilon}{2}\right) \mathbb{E}[\sum_{i: x_i^* = 0} \phi_i(v_i) \cdot x_i^*] \\
&= \left(1 - \frac{\epsilon}{2}\right) \mathbb{E}[\sum_i \phi_i(v_i) \cdot x_i^*] \qquad (5.17)
\end{aligned}$$

Where the first inequality follows from the fact that for any i such that $x_i > 0$ it holds $\phi_i(v_i) \geq \widetilde{\phi}_i(v_i)$, the second inequality follows from the fact that x is optimal w.r.t. the vector $\widetilde{\phi}(\boldsymbol{v})$, the first equality follows from the definition of the vector $\widetilde{\phi}(\boldsymbol{v})$ and the second one from the non-negativity of the vector x^*. The third inequality follows from the fact that for any i for which $x_i^* > 0$ as proven earlier it holds that $0 \leq \epsilon_i \leq \frac{\epsilon}{2} \cdot \phi_i(v_i)$, the third equality follows from linearity of expectations, the forth one from the fact that the second term is exactly 0, and the last one follows once again from linearity of expectations.
Equation 5.17 and the fact that the expected revenue of any mechanism is equal to its expected virtual social welfare immediately prove that the expected revenue of our extended polymatroid DA t-level auction is at least a $\left(1 - \frac{\epsilon}{2}\right)$-fraction of the optimal expected revenue. ☐

As we discussed earlier, for a fixed polymatroid constraint defined by some submodular function h on a set $\mathcal{N}$ bidders the maximum possible revenue of any mechanism is $H \cdot h(\mathcal{N})$. This observation, along with theorems 5.7, 5.8 and 5.1 immediately yield:

Corollary 5.4. *With probability at least $1-\delta$, for any environment with polymatroid constraints defined by some submodular function h, the empirical revenue maximizer for a sample set of size N of the class of extended polymatroid DA t-level auctions is a $1-O(\epsilon)$-fraction to Myerson for n linear bidders whose valuations per unit are in $[1, H]$, for $t = O\left(\frac{1}{\epsilon} + \log_{1+\epsilon} H\right)$ and*

$$N = O\left(\left(\frac{H \cdot h(S)}{\epsilon}\right)^2 \left(nt \log(nt) \ln \frac{H \cdot h(S)}{\epsilon} + \ln \frac{1}{\delta}\right)\right) = \widetilde{O}\left(\frac{H^2 h(S)^2 n}{\epsilon^3}\right),$$

By supplying polymatroid DA t-level auctions with those level scores we managed to reduce the number of levels required per bidder, for the same approximation to the optimal expected revenue, by about a factor of H, which in turn reduced the number of required samples, for the same parameters ϵ and δ by a factor of H.

Chapter 6

Learning Deferred-Acceptance Auctions for Multi-Parameter Environments

Now we will move to multi-parameter environments. For a multi-unit auction with m units available for sale, every bidder now declares m numbers, her *marginal values* for clinching an additional unit, provided she has already clinched 0 up to $m-1$ units. Once again, we will attempt to determine a mechanism class that, for a natural-enough set of assumptions about the bidders' distributions, given some samples of the bidders' valuations achieves high expected revenue. Multi-parameter environments introduce two new challenges: First, now the revenue-optimal mechanism is not known. As a result, we will have to compare the expected revenue of the mechanism class we propose directly against the expected social welfare of VCG, which is of course an upper bound on the expected revenue of any mechanism. Secondly, Secondly, for this kind of environments, the VCG mechanism is not implementable as a DA auction. In fact, even for 2 units available for sale, no DA auction can guarantee in a worst-case analysis an approximation ratio to the optimal social welfare that is better than $\frac{1}{\sqrt{2}}$. Since any dominant strategy incentive compatible mechanism has, for any valuation profile, revenue less or equal to its social welfare, this means that no DA auction can guarantee revenue that is a $\frac{1}{\sqrt{2}}$ fraction of the social welfare of VCG.

6.1 Distribution Assumptions

The Setting (Submodular Bidders, Ordered Standard Deviations)

- n bidders.
- multi-unit auction with m units available for sale.
- every bidder i declares m *marginal values* $v_{i,1}, v_{i,2}, \ldots, v_{i,m}$ to the mechanism, where $v_{i,j}$: Bidder i's value for acquiring her j-th unit of the good, provided she has already clinched $j-1$ units.
- every bidder's valuation function is submodular: $v_{i,1} \geq v_{i,2} \geq \cdots \geq v_{i,m}$ $\forall i$
- Every marginal value $v_{i,j}$ follows some Gaussian or Sub-Gaussian distribution $F_{i,j}$.

- For every bidder, the standard deviations of the distributions of her marginal values are ordered in the same way as those values: $\sigma_{i,1} \geq \sigma_{i,2} \geq \cdots \geq \sigma_{i,m}\ \forall i$

This set of assumptions is quite natural: A wide range of distributions are Sub-Gaussian, including Gaussian distributions and mixtures of Gaussian distributions, uniform distributions and all bounded distributions. Informally, any distribution with tails that are dominated by (i.e. decay as fast as) the tails of a Gaussian is Sub-Gaussian. Additionally, in nature oftentimes distributions with higher expected values also have higher variances. For example, this is a much weaker assumption than assuming that the standard deviations or variances of the distributions are proportional to their expected values. Finally, submodular valuation functions arise naturally in many environments where bidders exhibit "diminishing returns" for multiple units of the same good.

6.2 The Unit Bundling Mechanism Class

The mechanism class we propose for this setting has a very simple algorithmic description: For every bidder, the optimal reserve price and bundle size were predetermined in the samples. In the auction every bidder simply faces her predetermined reserve price. If her bid is higher than that, then she is allocated that number of units and she pays her reserve price. Else, she is finalized without clinching any units.

Definition 6.1 (Unit Bundling Mechanism Class - $\mathcal{UB}$)**.** *Every unit-bundling DA auction for n bidders and m units available for sale can be described by its parameters* $\{(s_1, s_2, \ldots, s_n) \in \mathbb{N}_+^n \mid s_1+s_2+\cdots+s_n \leq m\}$ *and* $r_1, r_2, \ldots, r_n \in \mathbb{R}_+$*. This set of* $2n$ *numbers define a unit bundling DA auction with the following allocation and payment rules:*

- *Offer to each bidder* i *a bundle of* s_i *units at a price of* r_i *per unit, so* $s_i \cdot r_i$ *in total.*
- *Bidders that accept their offer get the corresponding bundles at those prices, while the rest of the bidders are allocated no units and pay nothing.*

This mechanism class may look overly simplistic, but in reality this is not the case. Intuitively, DA auctions work well for bidders with linear valuation functions, but not for submodular ones. With this mechanism class the aim is to determine, based on the samples, for every bidder the optimal point in which to "straighten" her valuation function.
But first we need to prove that the auctions we just described are indeed valid generalized deferred acceptance auctions:

Proposition 6.1. *Any unit bundling DA auction, as described in definition 6.1, can be implemented as a valid generalized multi-parameter deferred-acceptance auction.*

Proof. Let $s_1, s_2, \ldots, s_n$ and $r_1, r_2, \ldots, r_n$ be the fixed parameters of some unit bundling auction, in accordance with definition 6.1. For every bidder i, we introduce in the original set of active bidders, A_1, a "duplicate" bidder i'. This duplicate bidder bids $s_i \cdot r_i$ and loses, according to the tie-breaking rule against bidder i, i.e. it holds $i' \prec i$. The sets of scoring and clinching functions of this multi-parameter deferred acceptance auctions are:

- Scoring function: $\sigma_i^{A_t}(b_i, b_{\mathcal{N}\setminus A_t}) = \sum_{j=1}^{s_i} b_{i,j}$
  where $b_{i,j}$: Bidder i's j-th declared marginal value

- $g_i^{A_t}(b_{\mathcal{N}\setminus A_t}) = \begin{cases} 0, & \text{if } i' \in A_t \\ s_i, & \text{else} \end{cases}$

For those scoring and clinching functions, the outcome of the auction will be exactly the one described in definition 6.1: If some bidder i bids less than $s_i \cdot r_i$ for i units, then she will have a lower score than her duplicate i', whose score will be exactly $s_i \cdot r_i$. Thus, she will be finalized before bidder i' and according to her clinching function, she will be allocated 0 units. Alternatively, if she bids at least $s_i \cdot r_i$, then she will be finalized after bidder i'. In this case, according to her clinching function she will be allocated s_i units. The price she will pay, according to Myerson payments, is the minimum she could have bid and still won those units which, according to the tie-breaking rule that favors i against i', is exactly $s_i \cdot r_i$.
The scoring function of every bidder is weakly increasing in her bid and does not depend on the bids of the other still-active bidders. So, it is a valid scoring function. The clinching function of every bidder is weakly increasing as the set of active bidders shrinks, so it is a valid clinching function. Thus, for any valid parameter vectors $\boldsymbol{s}$ and $\boldsymbol{r}$, the resulting auction can be implemented as a valid multi-parameter generalized DA auction. □

Our approach will be similar to the one we used in single-parameter environments: For the proposed mechanism class, we will establish generalization and representation error bounds. The main difference is that now in the representation bound, instead of the optimal expected revenue, we will have to compare the expected revenue of the optimal mechanism from the proposed mechanism class directly against the expected social welfare of VCG.

6.3 Generalization Error Bound

Our approach to bounding the generalization error of the unit bundling class is similar to what we did for polymatroid environments: First we will bound the pseudo-dimension of the class using the sample complexity framework introduced by Balcan in [12] and then we will apply Pollard's uniform convergence theorem.

Theorem 6.1. *The class of unit bundling auctions for n bidders and m units available for sale is $(2n, nm)$-delineable.*

Proof. For the parameter space $\mathcal{P}$ of the mechanism class we have:

- $(r_1, r_2, \ldots, r_n) \in \mathbb{R}_+^n$
- $(s_1, s_2, \ldots, s_n) \in \mathbb{N}^n$
- $s_1 + s_2 + \cdots + s_n \leq m$.

Fix a valuation profile $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$. Consider some fixed bidder i, and let $v_i = (v_{i,1}, v_{i,2}, \ldots, v_{i,m})$ be bidder i's marginal values in that valuation profile. With respect to bidder i, the parameter space $\mathcal{P}$ consists of $m+1$ connected components: Each connected component has some fixed value for the bundle size $s_i \in \{0, 1, \ldots, m\}$, while r_i can take any value in $\mathbb{R}_+$. Because s_i takes discrete values, regions of $\mathcal{P}$ with different values for s_i correspond to different connected components. Within each connected component the bundle size of bidder i is fixed and as the reserve price per unit is raised, the revenue from bidder i is raised, up to the point where the bidder is no longer interested in the bundle, and the revenue drops down to zero. Formally, let a be the value of the parameter s_i in some fixed connected component. Then, in that connected component the revenue from bidder i is:

$$\text{Rev}_{i,p}(\boldsymbol{v}) = \begin{cases} a \cdot r_i, \text{ if } \sum_{j=1}^{a} v_{i,j} \geq a \cdot r_i \\ 0, \text{ else} \end{cases} \tag{6.1}$$

It suffices to observe that in any connected component of $\mathcal{P}$, $s_i \cdot r_i$ is linear in its 2 parameters because in any connected component, s_i is fixed to some constant a. Thus, in any such connected component, there is only 1 hyperplane splitting it into 2 regions where the profit is linear: $\sum_{j=1}^{a} v_{i,j} \geq a \cdot r_i$. Because there are $m+1$ possible values for s_i, there are in total $m+1$ such connected components for bidder i, and m corresponding hyperplanes (for $s_i = 0$, the revenue from bidder i remains 0 regardless of r_i). Because $\mathcal{P}$ can only lead to feasible allocations and using the fact that the total revenue in any valuation profile is simply the sum of the revenues from every bidder, for every valuation profile $\boldsymbol{v}$, there are in total at most $n \cdot m$ hyperplanes splitting $\mathcal{P}$ in connected components, in each of whom the revenue of the mechanism class is linear in $\mathcal{P}$. □

Combining the above theorem with Balcan's theorem 4.2 immediately yields:

Theorem 6.2. *The pseudo-dimension of the class of unit bundling auctions is $O(n \log(nm))$.*

6.4 Representation Error Bound

In this section we will compare the expected revenue of the optimal mechanism from the unit bundling mechanism class against the expected social welfare of VCG. The main idea is to use an interface between these 2 mechanisms: a new mechanism, which we call *the a priori optimal mechanism*, $\mathcal{A}$. For this mechanism we can bound the difference between the expected social welfare of VCG and its expected social welfare, and we can also bound the difference between its expected social welfare and the expected social welfare of the optimal mechanism from the unit bundling class. The sum of these two differences is an upper bound on the difference between the expected social welfare of VCG and the expected revenue of the optimal mechanism from the unit bundling class, which constitutes a representation error bound.

Definition 6.2 (The A Priori Optimal Mechanism, $\mathcal{A}$)**.** *$\mathcal{A}$ is a hypothetical mechanism. It has perfect distribution knowledge and chooses, without taking the bids into consideration, the fixed allocation with the highest expected social welfare.*

In this setting $\mathcal{A}$ has a simple algorithmic description:

Algorithm 3: A Priori Optimal Mechanism for multi-unit auctions

Result: The fixed allocation with the highest expected social welfare
1. Sort all marginal values according to their expected value ;
2. Allocate to every bidder as many units as her expected marginal values in the m first positions of that list, breaking ties according to $\prec$;

Theorem 6.3. *For any distribution F over the bidders' valuation functions, the mechanism class of unit bundling auctions contains a mechanism $\mathcal{M}$ with expected revenue satisfying:*

$$\mathbb{E}_{\boldsymbol{v}}[Rev(\mathcal{M}, \boldsymbol{v})] \geq \mathbb{E}_{\boldsymbol{v}}[SW(\mathcal{A}, \boldsymbol{v})] - \sum_{i=1}^{n} (3\widehat{B}_i)^{1/3} \widehat{\sigma_i}^{2/3} \tag{6.2}$$

where $\widehat{B_i}$*: Bidder i's expected value for the number of units she is allocated by* $\mathcal{A}$*,*
$\widehat{\sigma_i}^2$*: The variance of bidder i's cumulative distribution for that number of units, i.e.*
$\widehat{\sigma_i}^2 = \sum_{j=1}^{s_i^*} \sigma_{i,j}^2$

For the proof we will need a lesser known result from [25].

Corollary 6.1 ([25], Corollary 2)**.** *The following bound holds for any distribution D:*

$$S \le (3B)^{1/3}\sigma^{2/3} \tag{6.3}$$

where B is the expected value of the distribution, σ^2 *is its variance and S is its separation.*

Proof. For any distribution F over the bidders' valuation functions, the unit bundling class contains a mechanism $\mathcal{M}$ that for every bidder i has the same bundle size as what that bidder was allocated in $\mathcal{A}$ and in addition, it has the optimal reserve price for that bundle and bidder. Using linearity of expectations, the expected revenue of $\mathcal{M}$ is simply the sum of its expected revenue on every bidder:

$$\mathbb{E}_{\boldsymbol{v}}[\text{Rev}(\mathcal{M}, \boldsymbol{v})] = \mathbb{E}_{\boldsymbol{v}}[\sum_{i=1}^{n} \text{Rev}_i(\mathcal{M}, \boldsymbol{v})] = \sum_{i=1}^{n} \mathbb{E}_{\boldsymbol{v}}[\text{Rev}_i(\mathcal{M}, \boldsymbol{v})] \tag{6.4}$$

where $\text{Rev}_i(\mathcal{M}, \boldsymbol{v})$: The revenue of $\mathcal{M}$ from bidder i on the valuation profile $\boldsymbol{v}$.
We will restrict our attention to some fixed bidder i. Let s_i^* be the amount of units she is allocated in $\mathcal{A}$ and therefore her bundle size in $\mathcal{M}$. Against bidder i, $\mathcal{M}$ faces a posted price problem. In posted price problems with a single bidder, we can view the expected revenue of the mechanism as the expected value of the bidder for the good (or bundle) available for sale, minus her expected *separation*, i.e. the part of the bidder's value for the good that the mechanism could not extract as revenue. Formally:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{v}}[\text{Rev}_i(\mathcal{M}, \boldsymbol{v})] &= \mathbb{E}_{\boldsymbol{v}}[\sum_{j=1}^{s_i^*} v_{i,j}] - S_i \\ &= \widehat{B}_i - \widehat{S}_i \end{aligned} \tag{6.5}$$

where $\widehat{B}_i$ is bidder i's expected value for a bundle of s_i^* units and $\widehat{S}_i$ is the the expected separation between her value for those units and the expected revenue.
Applying theorem 6.1 to the previous equation immediately yields:

$$\mathbb{E}_{\boldsymbol{v}}[\text{Rev}_i(\mathcal{M}, \boldsymbol{v})] \ge \widehat{B}_i - (3\widehat{B}_i)^{1/3}\widehat{\sigma_i}^{2/3} \tag{6.6}$$

Combining equations 6.4 and 6.6:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{v}}[\text{Rev}(\mathcal{M}, \boldsymbol{v})] &\ge \sum_{i=1}^{n} \left[\widehat{B}_i - (3\widehat{B}_i)^{1/3}\widehat{\sigma_i}^{2/3}\right] \\ &= \sum_{i=1}^{n} \widehat{B}_i - \sum_{i=1}^{n} (3\widehat{B}_i)^{1/3}\widehat{\sigma_i}^{2/3} \\ &= \mathbb{E}_{\boldsymbol{v}}[\text{SW}(\mathcal{A}, \boldsymbol{v})] - \sum_{i=1}^{n} (3\widehat{B}_i)^{1/3}\widehat{\sigma_i}^{2/3} \end{aligned} \tag{6.7}$$

Where the first equality follows from linearity of expectations and the second one from the fact that the expected social welfare of $\mathcal{A}$ is the sum of the expected value of all bidders in its allocation. □

The above theorem constitutes our bound between the expected revenue of the optimal mechanism from the unit bundling class and the expected social welfare of $\mathcal{A}$. As we discussed in the start of this section, the next step is to bound the difference between the expected social welfare of VCG and the expected social welfare of $\mathcal{A}$.

Theorem 6.4. *In the setting described at the start of section 6.1, for the expected social welfare of the a priori optimal mechanism, $\mathcal{A}$ it holds:*

$$\mathbb{E}_{\boldsymbol{v}}[SW(\mathcal{A},\boldsymbol{v})] \geq \mathbb{E}_{\boldsymbol{v}}[SW(\mathit{VCG},\boldsymbol{v})] - \sum_{i=1}^{m} \max\{\mu_{m+i} - \mu_{m+1-i} + \sigma_{m+i}\sqrt{2\log((n-1)m)} + \sigma_1\sqrt{2\log m}, 0\} \tag{6.8}$$

For this proof we will need one additional lemma from probability theory.

Lemma 6.1 ([26]). *Let $Y = \max_{1\leq i\leq n} X_i$, where $X_i \sim \mathcal{N}(\mu,\sigma^2)$. Then $\mathbb{E}[Y] \leq \mu + \sigma\sqrt{2\log n}$.*

Lemma Proof.

$$\begin{aligned}
\exp(t\mathbb{E}[Y]) &\leq \mathbb{E}[\exp(tY)] \\
&= \mathbb{E}[\max \exp(tX_i)] \\
&\leq \sum_{i=1}^{n} \mathbb{E}[\exp(tX_i)] \\
&= n\exp(t\mu + t^2\sigma^2/2)
\end{aligned}$$

The first inequality follows from Jensen's inequality, and the second one is the Union Bound. The last equality follows from the definition of the moment generating function.
Taking the logarithm of both sides of this inequality, we get

$$\mathbb{E}[Y] \leq \mu + \frac{\log n}{t} + \frac{t\sigma^2}{2}$$

This can be minimized by taking $t = \frac{\sqrt{2\log n}}{\sigma}$, which gives the desired result

$$\mathbb{E}[Y] \leq \mu + \sigma\sqrt{2\log n} \tag{6.9}$$

□

Proof. For any valuation profile $\boldsymbol{v}$, the allocation of $\mathcal{A}$ can be converted to the allocation of VCG using at most m *unit moves*, where a "move" means removing a unit from some bidder i' that was allocated more units by $\mathcal{A}$ than by VCG and allocating it to a bidder i' that was allocated less. Of course, since for a fixed distribution F over the bidders' valuations the allocation of $\mathcal{A}$ is fixed but the allocation of VCG is not, those moves depend on the actual valuation profile $\boldsymbol{v}$.
For any distribution F, the difference between the expected social welfare of VCG and the expected social welfare of $\mathcal{A}$, using linearity of expectations, is exactly the sum of the expected gain in social welfare by each of those m moves. We define an ordering on those moves:
move(i): For a given valuation profile $\boldsymbol{v}$, remove the unit that corresponds to the i-th lowest marginal value, out of the ones that $\mathcal{A}$ satisfied, but VCG did not (i.e. remove it from the corresponding bidder), and allocate it where it corresponds to the i-th highest *expected* marginal value out of the ones that VCG did satisfy and $\mathcal{A}$ did not.

Let gain(move(i)) be the gain in social welfare from move(i). It is obvious from the definition of those moves that:

$$\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathrm{VCG}, \boldsymbol{v}) - \mathrm{SW}(\mathcal{A}, \boldsymbol{v})] = \sum_{i=1}^{m} \mathbb{E}_{\boldsymbol{v}}[\mathrm{gain}(\mathrm{move(i)})] \tag{6.10}$$

Now it suffices to bound the quantity $\mathbb{E}_{\boldsymbol{v}}[\mathrm{gain}(\mathrm{move(i)})]$. This is quite an involved task. Let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{mn}$ be the expected values of all mn marginal values, in decreasing order and let $\overline{\sigma}_1, \overline{\sigma}_2, \ldots, \overline{\sigma}_{mn}$ be their corresponding standard deviations. In our setting, for the distributions of any bidder it holds that their standard deviations are ordered in the same way as their expected values, thus: $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_{mn}$.

We will restrict our attention to one such move, move(i) for some fixed i. For this move the lowest possible expected value of the distribution of the unit we are removing from the allocation of $\mathcal{A}$ is μ_{m+1-i} because by the definition of move(i), for the units that correspond to lower expected marginal values in the allocation of $\mathcal{A}$, if they were to be removed, they would have been removed by previous moves. The corresponding variance of the distribution of that unit is at most $\overline{\sigma}^2_{m+1-i}$.

Similarly, the highest possible expected value *of the distribution* of the unit we are adding, from the allocation of VCG that was not in the allocation of $\mathcal{A}$ is μ_{m+i} because again, by the definition of move(i), if units were to be allocated to marginal values with expected value in $\{\mu_{m+1}, \mu_{m+2}, \ldots, \mu_{m+i-1}\}$, they would have been allocated by previous moves. The corresponding variance of the distribution of that unit is at most $\overline{\sigma}^2_{m+i}$.

At a first glance, someone following along up to this point could assume that for the expected gain of the i-th move it holds: $\mathbb{E}_{\boldsymbol{v}}[\mathrm{gain}(\mathrm{move(i)})] \leq \mu_{m+i} - \mu_{m+1-i}$. However, this is not the case. This would be true only if the units removed and added by that move were two independent samples from their respective distributions. In reality, we have the prior knowledge that those units were chosen and not chosen by VCG respectively, and our expectation needs to account for that. The rest of the proof does exacty that.

When move(i) is allocating a unit to some marginal value, it only has $(n-1)m$ options, because m out of the nm marginal values are already included in the allocation of $\mathcal{A}$. Because both the mean and standard deviation/Sub-Gaussian parameter decrease as we "move down" in the two sorted lists:

- If all the $(n-1)m$ marginal values not included in the allocation of $\mathcal{A}$ were sampled from Gaussian/Sub-Gaussian distributions with parameters $(\mu_{m+i}, \sigma^2_{m+i})$ (the highest possible), their expected maximum would be at most $\mu_{m+i} + \overline{\sigma}_{m+i}\sqrt{2\log((n-1)m)}$.
- If any of the $(n-1)m$ marginal values follows a distribution with either smaller mean or standard deviation, the expected value of their maximum strongly decreases.
- Thus, using lemma 6.1 the expected value of the unit that the i-th move is adding to the allocation of $\mathcal{A}$ is at most $\mu_{m+i} + \sigma_{m+i}\sqrt{2\log((n-1)m)}$.

In the same vein, when move(i) is removing a unit from the allocation of $\mathcal{A}$, it only has m options, because $\mathcal{A}$ allocates at most m units. Because both the mean and standard deviation/Sub-Gaussian parameter decrease as we "descend" the two sorted lists:

- The maximum possible variance of the distribution of a unit in the allocation of $\mathcal{A}$ is σ_1^2.
- If all the m marginal values in the allocation of $\mathcal{A}$ were sampled from Gaussian/Sub-Gaussian distributions with parameters $(\mu_{m+1-i}, \sigma_1^2)$, their expected minimum would be at least $\mu_{m+1-i} - \sigma_1\sqrt{2\log(m)}$.

- If any of the m marginal values follows a distribution with greater mean or lower variance, the expected value of their minimum strongly increases.
- Thus, using lemma 6.1 the expected value of the unit that the i-th move is removing from the allocation of $\mathcal{A}$ is at least $\mu_{m+1-i} - \sigma_1\sqrt{2\log(m)}$.

Since the expected gain of every move, if performed, is simply the expected difference between the unit being added and the unit being removed from the allocation of $\mathcal{A}$:

$$\mathbb{E}_{\boldsymbol{v}}[gain(move(i))] \leq \mu_{m+i} - \mu_{m+1-i} + \sigma_{m+i}\sqrt{2\log((n-1)m)} + \sigma_1\sqrt{2\log m} \tag{6.11}$$

Combining equations 6.11 and 6.10 with the fact that any move that will actually be performed results in non-negative gain in social welfare:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{v}}[\text{SW}(\text{VCG},\boldsymbol{v}) - \text{SW}(\mathcal{A},\boldsymbol{v})] &= \sum_{i=1}^{m}\mathbb{E}_{\boldsymbol{v}}[\text{gain}(\text{move(i)})]\\ &\leq \sum_{i=1}^{m}\max\{\mu_{m+i} - \mu_{m+1-i} + \overline{\sigma}_{m+i}\sqrt{2\log((n-1)m)} + \overline{\sigma}_1\sqrt{2\log m}, 0\}\end{aligned} \tag{6.12}$$

Using linearity of expectations and solving with respect to $\mathbb{E}_{\boldsymbol{v}}[\text{SW}(\mathcal{A},\boldsymbol{v})]$ immediately yields the desired bound. □

This result has a very intuitive interpretation: Notice that the difference $\mu_{m+i} - \mu_{m+1-i}$ is always non-positive. Holding the variances/Sub-Gaussian parameters of all the distributions fixed, as the "distance" between the expected values of the distributions increases, the difference between the expected social welfare of VCG and $\mathcal{A}$ decreases. Now that the distributions are "further apart" it is less likely for a marginal value to surpass some other with higher expected value, thus VCG has a smaller chance of finding suitable units to swap, different than those of $\mathcal{A}$. Conversely, holding the expected values of the distributions of all marginal values fixed, as their variances/Sub-Gaussian parameters increase, so does the difference between the expected social welfare of VCG and $\mathcal{A}$. For higher Sub-Gaussian parameters, the tails of the distributions decay slower, therefore the regions and the probability mass where distributions of different marginal valuations are overlapping increases. This means that VCG has a higher chance of finding suitable units to "swap", and therefore increase its social welfare compared to $\mathcal{A}$.
Combining theorems 6.2, 6.3 and 6.4 with Pollard's uniform convergence theorem and solving for the expected revenue results in the following corollary:

Corollary 6.2. *In the setting described at the start of section 6.1, for the empirically optimal mechanism $\mathcal{M} \in \mathcal{UB}$ on a sample set of N valuation profiles, with probability at least $1-\delta$ it holds:*

$$\begin{aligned}\mathbb{E}_{\boldsymbol{v}}[Rev(\mathcal{M},\boldsymbol{v})] \geq \mathbb{E}_{\boldsymbol{v}}[SW(VCG,\boldsymbol{v})] - \sum_{i=1}^{n}(3\widehat{B}_i)^{1/3}\widehat{\sigma}_i^{\,2/3} - O\left(U\sqrt{n\log(nm)/N} + U\sqrt{\ln(1/\delta)/N}\right)\\ - \sum_{i=1}^{m}\max\{\mu_{m+i} - \mu_{m+1-i} + \overline{\sigma}_{m+i}\sqrt{2\log((n-1)m)} + \overline{\sigma}_1\sqrt{2\log m}, 0\}\end{aligned} \tag{6.13}$$

where U: The maximum profit achievable by mechanisms in $\mathcal{UB}$.

6.5 Relaxing the Distribution Assumptions

In this section we will relax the assumptions about the bidders' valuation distributions and show that a result very similar to corollary 6.2 still holds. Now the only assumptions that we make are that all bidders have submodular valuation functions and their marginal values follow exclusively Gaussian and Sub-Gaussian distributions. Formally:

The Setting (Submodular Bidders, Sub-Gaussian Distributions)

- n bidders.
- multi-unit auction with m units available for sale.
- every bidder i declares m *marginal values* $v_{i,1}, v_{i,2}, \ldots, v_{i,m}$ to the mechanism, where $v_{i,j}$: Bidder i's value for acquiring her j-th unit of the good, provided she has already clinched $j-1$ units.
- every bidder's valuation function is submodular: $v_{i,1} \geq v_{i,2} \geq \cdots \geq v_{i,m}\ \forall i$
- Every marginal value $v_{i,j}$ follows some Gaussian or Sub-Gaussian distribution $F_{i,j}$.

As we alluded to when commenting on the previous setting (6.1), this set of assumptions is quite natural: A wide range of distributions are Sub-Gaussian, including Gaussian distributions and mixtures of Gaussian distributions, uniform distributions and all bounded distributions. Furthermore, Gaussian distributions are perhaps the most common in nature, and it is a common practice in statistical theory to use Gaussian distributions to represent real-valued distributions whose distributions are unknown ([27, 28]). Finally, submodular valuation functions arise naturally in many environments where bidders exhibit "diminishing returns" for multiple units of the same good.

Theorem 6.5 (Alternative to 6.4)**.** *In the setting describe above, for the expected social welfare of the a priori optimal mechanism, $\mathcal{A}$, it holds:*

$$\mathbb{E}_{\boldsymbol{v}}[SW(\mathcal{A}, \boldsymbol{v})] \geq \mathbb{E}_{\boldsymbol{v}}[SW(VCG, \boldsymbol{v})] - \sum_{i=1}^{m} \overline{\sigma}_i \sqrt{2\log(nm)}, \tag{6.14}$$

where $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_{mn}$ is a total ordering on the standard deviations of the distributions of all nm marginal values.

Proof. The m marginal values that $\mathcal{A}$ satisfies are sampled from Gaussian/Sub-Gaussian distributions with parameters $(\mu_1, \sigma_1^2), (\mu_2, \sigma_2^2), \ldots, (\mu_m, \sigma_m^2)$ respectively. Thus, the expected social welfare of the allocation of $\mathcal{A}$ is exactly $\sum_{i=1}^{m} \mu_i$. This is greater or equal the a priori expected social welfare of any other allocation.
Let $\boldsymbol{v}$ be a random sample of the players' values, in which we compare the social welfare of $\mathcal{A}$ with that of VCG. The expected social welfare of $\mathcal{A}$ on that sample, as on any other sample, is $\sum_{i=1}^{m} \mu_i$. The expected social welfare of VCG is a bit harder to analyze:

- In any valuation profile $\boldsymbol{v}$, VCG will allocate units in such a way that they correspond to the m out of the nm highest marginal values.
- Suppose that on that valuation profile $\boldsymbol{v}$, VCG allocated a unit that corresponded to some marginal value i. Let (μ_i, σ_i^2) be the expected value and variance/Sub-Gaussian parameter of the distribution of that marginal value i.

- There are nm marginal values in total. Even if all of them were i.i.d. samples from the distribution of i, the expected value of the maximum of those nm samples, which is the optimal that VCG could have picked, is at most $\mu_i + \sigma_i\sqrt{2\log(nm)}$ (Application of Lemma 6.1).
- Thus, the expected value of some marginal value i on a sample, provided that VCG picked that marginal value on that sample, can only be at most $\sigma_i\sqrt{2log(nm)}$ greater than the a priori expected value of that same marginal value, where σ_i: The standard deviation/Sub-Gaussian parameter of the distribution of i.
- Using the fact that VCG, just like any other mechanism, can allocate at most m units, summing up over all the units allocated by the VCG mechanism, their expected value can be at most $\sum_{i=1}^{m}\overline{\sigma_i}\sqrt{2\log(nm)}$ greater than their a priori expected value.

That, combined with the fact that $\mathcal{A}$ has a priori expected social welfare greater or equal to that of any other allocation, immediately implies:

$$\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathcal{A},\boldsymbol{v})] \geq \mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathrm{VCG},\boldsymbol{v})] - \sum_{i=1}^{m}\overline{\sigma}_i\sqrt{2\log(nm)} \tag{6.15}$$

□

Combining theorems 6.2, 6.3 and 6.5 (instead of 6.4) with Pollard's uniform convergence theorem and solving for the expected revenue yields:

Corollary 6.3. *In the setting described at the start of this section, for the empirically optimal mechanism $\mathcal{M} \in \mathcal{UB}$ on a sample set of N valuation profiles, with probability at least $1-\delta$ it holds:*

$$\begin{aligned}\mathbb{E}_{\boldsymbol{v}}[Rev(\mathcal{M},\boldsymbol{v})] \geq \mathbb{E}_{\boldsymbol{v}}[SW(VCG,\boldsymbol{v})] - \sum_{i=1}^{n}(3\widehat{B}_i)^{1/3}\widehat{\sigma}_i^{\,2/3} - \sum_{i=1}^{m}\overline{\sigma}_i\sqrt{2\log(nm)}\\ - O\left(U\sqrt{n\log(nm)/N} + U\sqrt{\ln(1/\delta)/N}\right)\end{aligned} \tag{6.16}$$

where U is the maximum profit achievable by mechanisms in $\mathcal{UB}$, $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_{mn}$ is a total ordering on the standard deviations of the distributions of all nm marginal values and $\widetilde{B}_i, \widetilde{\sigma}_i^{\,2}$ are the expected value and variance of $\widetilde{F}_i$, bidder i's valuation distribution for x_i units, i.e. $\widetilde{\sigma}_i^{\,2} = \sum_{j=1}^{x_i}\sigma_{i,j}^2$.

At a first glance, it is hard to grasp if the above result is good. The expected revenue of $\mathcal{M}$ is greater or equal to the expected social welfare of VCG, which is an obvious upper bound on the expected revenue of any mechanism, minus three small, additive terms. The quantity $\sum_{i=1}^{n}(3\widehat{B}_i)^{1/3}\widehat{\sigma}_i^{\,2/3}$, if the third root over every term didn't exist and the total variances of the players were all equal to one would be equal to 3 times the expected social welfare of VCG. Of course, by applying a third root over each of its terms it becomes an order of magnitude smaller. The term $\sum_{i=1}^{m}\overline{\sigma}_i\sqrt{2\log(nm)}$ is sublinear in the number of bidders n and (almost) linear in the number of units available for sale, m. For high-valued goods, this term compared to the expected social welfare of VCG becomes trivially small. The final quantity, $O\left(U\sqrt{n\log(nm)/N} + U\sqrt{\ln(1/\delta)/N}\right)$, is the generalization error of the proposed mechanism class. Reasoning about this quantity is a bit harder, especially if one is not very familiar with sample complexity analysis. This term is a direct result of applying Pollard's Uniform Convergence Theorem (theorem 4.1), as is standard on learning theory, on the pseudo-dimension bound that we established for the unit bundling mechanism class. In our case, this bound depends on the number of bidders, n, and the number of units available for sale, m. But under non-anonymous prices, the dependence on the number of bidders is expected. Finally, when the valuations of the bidders are unit-dependent, the sample complexity bound should also depend on the number of units available for sale. For a more intuitive explanation behind the last 2 statements, we refer the interested reader to [12, sections 2 and 3].

Chapter 7

Restricting The Number of Samples

Our results up to this point were very promising. For any setting we studied, we showed that given enough samples, it is possible to learn a deferred-acceptance auction with high revenue. In some high-impact markets, like spectrum allocation where DA auctions are used in practice, auctions can be quite rare. For those environments, assuming a non-constant number of samples, or any other knowledge of the bidders' valuation distributions, might be unrealistic. Therefore, a logical next step for this thesis would be to design *prior-independent* auctions that, using a very small number of samples, achieve expected revenue that is a constant approximation of what can be achieved with an unrestricted number of samples. In this chapter we tackle exactly this problem, both for single and multi-parameter environments.

7.1 Single-Parameter Environments

The first single-parameter setting we will study is once again a multi-unit auction with bidders with linear valuations, as we studied in most of chapter 5. The only difference is that now the only requirement for the valuation distribution of every bidder is that it is regular, instead of having bounded support in some range $[1, H]$. This is a much weaker assumption. Formally:

The Setting (Budget-Additive Bidders, Regular distributions)

- n bidders.
- multi-unit auction with m units available for sale.
- every bidder i has a publicly known demand d_i.
- every bidder is additive, up to her demand d_i:
  For every bidder i, her value for acquiring x_i units of the good (a service level x_i), up to d_i, is $x_i \cdot v_i$, where v_i: her value per unit of the good. For more than d_i units: her value remains $v_i \cdot d_i$. Valuation functions of this form are called *budget-additive*.
- Every bidder i's value per unit, v_i, follows some distribution with probability density function $f_i(\cdot)$ and virtual valuation function $\phi_i(\cdot)$.
- Those distributions are regular.

Our approach for single-parameter environments was heavily inspired by Hartline and Roughgarden's "Simple versus Optimal Mechanisms". In our case however, some non-trivial modifications of their results were required.
In their paper, Hartline and Roughgarden introduced the notion of a *duplicated environment.* Intuitively, in a duplicated environment every bidder i is competing with every other bidder in the same way as she was in the original environment, but she is now competing with one additional bidder: her duplicate i'. Regardless of the exact details of the original environment, in the duplicated environment at most one of i, i' may have a non-zero allocation. Formally:

Definition 7.1 (Duplication of a Single-Parameter Environment, [29])**.** *Each bidder i with valuation distribution F_i is replaced by a pair i, i' whose valuations are i.i.d. draws from F_i. The feasible allocations in the duplicated environment are those satisfying:*

1. *At most one bidder from each pair is allocated any units.*
2. *The allocation, when naturally interpreted as an allocation in the original environment (i.e. all units that were allocated to either i or i' are instead allocated to i) is a feasible allocation in that environment.*

Another important notion that Hartline and Roughgarden introduced was that of *commensurate mechanisms.* They introduced that notion for binary environments, meaning that every bidder can either be a winning or a losing bidder. We extended that notion naturally to generalized, single-parameter environments and then proved that our extension maintains the desirable properties of the original definition.

Definition 7.2 (Single-Parameter Commensurate)**.** *Let $\mathcal{M}$ and $\mathcal{M}'$ be two mechanisms for a given single-parameter environment. Let $x_i(\boldsymbol{v})$ and $x'_i(\boldsymbol{v})$ denote the allocation of the two mechanisms to bidder i in the valuation profile $\boldsymbol{v}$. The mechanism $\mathcal{M}$ is single-parameter commensurate with $\mathcal{M}'$ if:*

$$\mathbb{E}_{\boldsymbol{v}}\left[\sum_{i:x_i(\boldsymbol{v})\neq x'_i(\boldsymbol{v})} x_i(\boldsymbol{v})\phi_i(v_i)\right] \geq 0 \tag{C1}$$

and

$$\mathbb{E}_{\boldsymbol{v}}\left[\sum_{i\in\mathcal{N}} p_i(\boldsymbol{v})\right] \geq \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i:x_i(\boldsymbol{v})\neq x'_i(\boldsymbol{v})} x'_i(\boldsymbol{v})\phi_i(v_i)\right] \tag{C2}$$

where $p_i(\boldsymbol{v})$: The payment of bidder i in $\mathcal{M}$ on valuation profile $\boldsymbol{v}$, for $x_i(\boldsymbol{v})$ units.

The formal definition of commensurate mechanisms may look complicated at first glance. The first condition (C1) requires that the expected virtual social welfare, but summed over only the bidders with different allocations by the two mechanisms, is non-negative for the allocation of $\mathcal{M}$. This assertion is generally non-trivial: Even though the unconditional expected value of a bidder's virtual valuation is zero, now there is the implicit conditioning on those bidders having different allocations in $\mathcal{M}'$. The second condition (C2) asserts that the expected total payments from bidders in $\mathcal{M}$ is greater than the expected virtual social welfare in $\mathcal{M}'$, but summed over only the bidders where the allocations of the two mechanisms differ.
At the heart of many of the results in [29] was the fact that satisfying the conditions of 7.2 is a sufficient condition for the expected revenue of the mechanism $\mathcal{M}$ to be a 2-approximation to that of the mechanism $\mathcal{M}'$.

Lemma 7.1 (Similar to [29], Lemma 3.9)**.** *If a mechanism $\mathcal{M}$ is single-parameter commensurate with a mechanism $\mathcal{M}'$, then*

$$\mathbb{E}_{\boldsymbol{v}}\left[Rev(\mathcal{M}, \boldsymbol{v})\right] \geq \frac{1}{2} \cdot \mathbb{E}_{\boldsymbol{v}}\left[Rev(\mathcal{M}', \boldsymbol{v})\right]. \tag{7.1}$$

Proof. We argue separately that

$$\mathbb{E}_{\boldsymbol{v}}\left[\mathrm{Rev}(\mathcal{M}, \boldsymbol{v})\right] \geq \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i: x_i(\boldsymbol{v}) = x'_i(\boldsymbol{v})} x'_i(\boldsymbol{v}) \phi_i(v_i)\right] \tag{7.2}$$

and

$$\mathbb{E}_{\boldsymbol{v}}\left[\mathrm{Rev}(\mathcal{M}, \boldsymbol{v})\right] \geq \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i: x_i(\boldsymbol{v}) \neq x'_i(\boldsymbol{v})} x'_i(\boldsymbol{v}) \phi_i(v_i)\right] \tag{7.3}$$

Adding these two and applying linearity of expectations and Myerson's lemma yields the theorem: The left-hand side equals $2 \cdot \mathbb{E}_{\boldsymbol{v}}\left[\mathrm{Rev}(\mathcal{M}, \boldsymbol{v})\right]$ and the right-hand side equals $\mathbb{E}_{\boldsymbol{v}}\left[\mathrm{Rev}(\mathcal{M}', \boldsymbol{v})\right]$. To derive inequality 7.2, write

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{v}}\left[\mathrm{Rev}(\mathcal{M}, \boldsymbol{v})\right] &= \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i \in \mathcal{N}} x_i(\boldsymbol{v}) \phi_i(v_i)\right] \\
&= \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i: x_i(\boldsymbol{v}) \neq x'_i(\boldsymbol{v})} x_i(\boldsymbol{v}) \phi_i(v_i)\right] + \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i: x_i(\boldsymbol{v}) = x'_i(\boldsymbol{v})} x_i(\boldsymbol{v}) \phi_i(v_i)\right] \\
&\geq \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i: x_i(\boldsymbol{v}) = x'_i(\boldsymbol{v})} x_i(\boldsymbol{v}) \phi_i(v_i)\right] = \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i: x_i(\boldsymbol{v}) = x'_i(\boldsymbol{v})} x'_i(\boldsymbol{v}) \phi_i(v_i)\right]
\end{aligned} \tag{7.4}$$

where the first equality follows from Myerson's lemma, the second from linearity of expectations and the inequality follows from condition (C1).
Deriving inequality 7.3 is straightforward:

$$\mathbb{E}_{\boldsymbol{v}}\left[\mathrm{Rev}(\mathcal{M}, \boldsymbol{v})\right] = \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i \in \mathcal{N}} p_i(\boldsymbol{v})\right] \geq \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i: x_i(\boldsymbol{v}) \neq x'_i(\boldsymbol{v})} x'_i(\boldsymbol{v}) \phi_i(v_i)\right] \tag{7.5}$$

where the equality is the definition of revenue and the inequality follows immediately from condition (C2). □

The reason we extended the definition of commensurate mechanisms, and the lemma regarding their expected revenue, from binary to arbitrary single-parameter environments is that it could be applied to multi-unit auctions with budget-additive bidders.

Lemma 7.2 (Similar to [29] Lemma 4.5)**.** *In a multi-unit auction with budget-additive bidders and regular valuation distributions, VCG with duplicates is single-parameter commensurate with the optimal mechanism, Myerson, without duplicates.*

Proof. We begin with the first requirement, (C1) of our definition. Let $\mathcal{N}'$ denote the set of all "duplicated" bidders and $\boldsymbol{v'}$ their valuation profile. Finally, let $x(\boldsymbol{v}, \boldsymbol{v'}), x'(\boldsymbol{v}, \boldsymbol{v'})$ denote the allocation of the VCG mechanism (with duplicates) and the optimal mechanism (without duplicates) respectively. By definition, $x'(\boldsymbol{v}, \boldsymbol{v'})$ is independent of $\boldsymbol{v'}$, and cannot allocate any items to duplicated bidders: $x'_{i'}(\boldsymbol{v}, \boldsymbol{v'}) = 0\ \forall i' \in \mathcal{N}'$.
Condition on $\boldsymbol{v}$ but not on $\boldsymbol{v}'$; this fixes $x'(\boldsymbol{v}, \boldsymbol{v}')$. We argue that

$$\mathbb{E}_{\boldsymbol{v'}}\left[\sum_{i\in\mathcal{N}\cup\mathcal{N}':\, x_i(\boldsymbol{v},\boldsymbol{v'})\neq x'_i(\boldsymbol{v},\boldsymbol{v'})} x_i(\boldsymbol{v}, \boldsymbol{v'})\phi_i(v_i)\right] \geq 0 \tag{7.6}$$

The unconditional inequality in (C1) follows. We prove (7.6) by showing that the expected combined contribution of each original bidder i and her duplicate i' to the left-hand side is non-negative. If one of i, i' has positive allocation, it is the bidder with higher valuation and hence, by regularity, with higher virtual valuation.
First consider an original bidder i such that $x'_i(\boldsymbol{v}, \boldsymbol{v}') > 0$. Since the valuation distributions are regular, the optimal mechanism only selects bidders with a non-negative virtual valuation, so $\phi_i(v_i) \geq 0$. It follows that the contribution from i, i' to the left-hand side of (7.6) in this case is non negative with probability 1: if $x_i(\boldsymbol{v}, \boldsymbol{v}') = x_{i'}(\boldsymbol{v}, \boldsymbol{v}') = 0$ the contribution to the virtual social welfare is zero; otherwise it is $\max\{x_{i'}(\boldsymbol{v}, \boldsymbol{v}'), x_i(\boldsymbol{v}, \boldsymbol{v}')\} \cdot \max\{\phi_i(v_i), \phi_i(v_{i'})\} \geq$
$\geq \max\{x_{i'}(\boldsymbol{v}, \boldsymbol{v}'), x_i(\boldsymbol{v}, \boldsymbol{v}')\} \cdot \phi_i(v_i) \geq 0$.
Now suppose that for the original bidder i we have $x'_i(\boldsymbol{v}, \boldsymbol{v}') = 0$. Condition further on the valuations $\boldsymbol{v}'_{-i}$ of all duplicates other than i', and let $\mathcal{E}$ denote the event that either $x_i(\boldsymbol{v}, \boldsymbol{v}') > 0$ or $x_{i'}(\boldsymbol{v}, \boldsymbol{v}') > 0$ (both of these cannot occur at the same time). If $\neg\mathcal{E}$ occurs, then the contribution from i, i' to the left-hand side of (7.6) is zero. Since $\boldsymbol{v}, \boldsymbol{v'}_{-i}$ are fixed, event $\mathcal{E}$ occurs if and only if $v_{i'}$ is at least some non-negative threshold t. In this case, the expected contribution of i, i' is $\mathbb{E}_{v_{i'}}[\max\{x_i(\boldsymbol{v}, \boldsymbol{v}')\phi_i(v_i), x_{i'}(\boldsymbol{v}, \boldsymbol{v}')\phi_i(v_{i'})\}]$, conditioned on $\boldsymbol{v}, \boldsymbol{v}'_{-i}$ and $\mathcal{E}$. This is lower bounded by the analogous conditional expectation of $x_{i'}(\boldsymbol{v}, \boldsymbol{v}')\phi_i(v_{i'})$, which is equivalent to

$$E_{v_{i'}}[x_{i'}(\boldsymbol{v}, \boldsymbol{v}')\phi_i(v_{i'})|v_{i'} \geq t]. \tag{7.7}$$

Since the unconditional expectation of a virtual valuation is zero, ϕ_i is non-decreasing by regularity and $x_{i'}(\boldsymbol{v}, \boldsymbol{v}') \geq 0$, the quantity in (7.7) is non-negative. Taking expectations over whether or not $\mathcal{E}$ occurs, and then over $\boldsymbol{v}'_{-i}$ completes the argument.
For the proof of condition (C2), we will have to improvise. We will prove the condition pointwise, for each valuation profile $\boldsymbol{v}, \boldsymbol{v}'$. For any valuation profile, VCG in the duplicated environment allocates at least as many units as the optimal mechanism in the original environment, since the allocation of the original environment is always feasible in the duplicated environment, and VCG chooses a maximal solutions with respect to social welfare (and no negative values exist). Let i be the bidder with the highest virtual value per unit $\phi_i(v_i)$ to whom the optimal mechanism in the original environment allocated more units than VCG with reserves did, in the duplicated environment, i.e. $x'_i(\boldsymbol{v}, \boldsymbol{v}') > x_i(\boldsymbol{v}, \boldsymbol{v}')$. Since $x'_i(\boldsymbol{v}, \boldsymbol{v}') > x_i(\boldsymbol{v}, \boldsymbol{v}')$, allocating an additional unit to bidder i was also possible on the duplicated environment (the demands of all pairs of original/duplicated bidders are the same). Thus, for any unit that was allocated differently by the VCG with reserves, the externality that that allocation caused was at least v_i. By the definition of VCG, for every unit allocated differently, for the payment p that the winner of that unit has to make it holds $p \geq v_i \geq \phi_i(v_i)$,
where $v_i \geq \phi_i(v_i)$ follows by the definition of a virtual valuation. Summing up over all units allocated differently by VCG with duplicates, and using the fact that it always allocates at

least as many units as the optimal mechanism in the original environment, we have that the total payments procured for those units by VCG with reserves is greater or equal the sum of virtual values collected by the optimal mechanism, for all units that it allocated differently than VCG with duplicates. This, combined with the fact that for the units that were allocated in the same way both mechanisms earn payments equal to the corresponding (same) virtual value completes the proof. □

The reason that we are interested in VCG with duplicates is that for a multi-unit auction, as for any single-parameter environment with budget-additive bidders, VCG with duplicates is implementable as a valid generalized DA auction.

Proposition 7.1. *In any environment with polymatroid constraints, including a multi-unit auction, VCG with duplicates is implementable as a generalized deferred-acceptance auction.*

Proof. The proof is similar to [18, Section 4]. The main difference is that now the allocation of every bidder has to adhere not only to the polymatroid constraint of the original environment, but also to the constraint that for every pair of an original bidder and her duplicate, at most one of them can have a non-zero allocation.
Let $h : 2^n \to \mathbb{R}_+$ be the submodular function defining the set of feasible outcomes:

$$P_h = \left\{ l \in \mathbb{N}^n \mid \sum_{i \in S} x_i \leq h(S) \; \forall S \subset N \right\} \tag{7.8}$$

In any problem instance involving polymatroid constraints, there exists a simple generalized DA auction that achieves the optimal social welfare in the *duplicated environment*. Given the submodular function h of the polymatroid constraint, the scoring and clinching functions of this auction are defined as follows:

- The polymatroid auction scoring function is

$$\sigma_i^{A_t}(b_i, b_{N \setminus A_t}) = b_i \tag{7.9}$$

- The polymatroid auction clinching function is

$$g_i^{A_t}(b_{N \setminus A_t}) = \begin{cases} h(A_t) - h(A_t \setminus \{i\}), & \text{if } \widetilde{i} \notin A_t \\ 0, & \text{else} \end{cases} \tag{7.10}$$

where with $\widetilde{i}$ we denote for any bidder i of the duplicated environment, the other bidder of the pair of original and duplicated bidder that i belongs to.
The scoring function of every bidder i is obviously valid, since it is weakly increasing in the bid of i and does not depend on the bids of the other still-active bidders. The clinching function is also valid: At each stage t each bidder $i \in A_t$ has clinched an allocation equal to her marginal contribution to the value of $h(A_t)$. Since h is submodular, this marginal contribution weakly increases as A_t shrinks, so this is a valid clinching function. Thus, the mechanism proposed is a valid single-parameter generalized deferred-acceptance auction in compliance with definition 4.11.
One can easily verify that this auction always yields the maximum possible social welfare. Note that its outcome is exactly the same as the one that would arise if we instead used the following forward-greedy algorithm: First give the bidder i with the highest valuation the highest allocation possible, i.e. $h(\{i\})$ (note: as in the definition of duplicated environments,

in the function h the units allocated to duplicated bidders are naturally interpreted as given to the original ones, e.g. $h(\{i'\}) = h(\{i\})$). Then, give the second highest bidder j the highest allocation possible, given the existing assignment to i, i.e. j gets $h(\{i, j\}) - h(\{i\})$ if her duplicate has not already been allocated any units, else she gets 0, and so on. This greedy algorithm has been proven to be optimal for polymatroid settings ([30]).

□

Combining lemmata 7.1 and 7.2 with proposition 7.1 immediately yields:

Theorem 7.1. *In the setting described at the start of section 7.1, there exists a simple generalized single-parameter auction that achieves, in the duplicated environment, at least half the expected revenue of the optimal mechanism, Myerson, in the original environment.*

The above theorem required quite a bit of work to be proven. One may wonder what was the point behind it; after all, the mechanism designer cannot duplicate the bidders of an auction. What the mechanism designer can do however is, if given a sample from every bidder's valuation distribution, use that sample to "simulate" the duplication of the environment.

Corollary 7.1. *For every multi-unit auction setting with bidders with linear valuations drawn independently from distributions that satisfy the regularity condition, there exists a generalized single-parameter deferred-acceptance auction that, using a single sample, achieves expected revenue that is a $\frac{1}{4}$-fraction of the expected revenue of the optimal auction.*

Proof. By theorem 7.1, there exists a generalized DA auction that achieves, in the duplicated environment at least half the expected revenue of the optimal auction in the original environment. In the duplicated environment, symmetry dictates that for every pair (i, i') of an original bidder and her duplicate, if one of them has non-negative allocation, then it is equally likely to be the original bidder or her duplicate. Thus, the expected revenue from the original bidders is at least a quarter of that of the optimal auction in the original environment.
Given a sample $\boldsymbol{s} = (s_1, s_2, \dots, s_n)$ of every bidder's valuation distribution, we can use that sample to simulate the duplication of the original environment: For every bidder i with demand d_i and valuation v_i her duplicate i' also has demand d_i and valuation s_i per unit. Then, we can run the auction described in proposition 7.1 on this extended set of bidders. Because we have restricted our attention to multi-unit auctions, the polymatroid DA auction described in that proposition takes a special form. Let b_i be the bid of bidder i in the auction, whether she is an original or a simulated bidder. Then, the scoring and clinching functions of the auction are:

- The multi-unit scoring function is

$$\sigma_i^{A_t}(b_i, b_{N \setminus A_t}) = b_i \tag{7.11}$$

- The multi-unit clinching function is

$$g_i^{A_t}(b_{N \setminus A_t}) = \min\left\{ d_i, \max\left\{ 0, m - \sum_{j \in A_t} \frac{d_j}{|\{j, \widetilde{j}\} \cap A_t|} \right\} \right\} \tag{7.12}$$

where $\widetilde{j}$ is the other bidder that belongs in the same pair of original, duplicate as j.

□

There is one interesting remark regarding the term $d_j/|\{j,\widetilde{j}\} \cap A_t|$. When every bidder is finalized, the aim is that she is allocated up to her demand as many units as possible, provided that the units left afterwards are enough to satisfy the demands of the still-active bidders. By dividing by the demand of every still-active bidder by $|\{j,\widetilde{j}\} \cap A_t|$ we are ensuring that in the case that both an original bidder and her duplicate are still active, then for that pair of bidders their demand will only be accounted for once. The reasoning is that from every such pair of bidders, in any feasible allocation at most one of them is allocated any units.

One intuitive way of interpreting this mechanism is the following: In our sample complexity approach, the proposed mechanism was a modified version of t-level auctions, where the levels were used as a discrete approximation of every bidder's virtual valuation function. Now that for every bidder only a single sample of her valuation distribution is available, we can field a modified version of t-level auctions with a single level and threshold per bidder: Whether a bidder is at level 0 or -1 corresponds to whether or not she has surpassed her duplicate. For bidders that are at level 0, ties between them are broken according to their bids.

7.1.1 Generalizing to Environments with Polymatroid Constraints

In our sample complexity approach, we managed to extend our results for multi-unit auctions to environments with polymatroid constraints. A question that arises naturally is if the same generalization can be achieved for our single-sample result. The answer to that question is positive. In fact, our approach will be the same as it was for multi-unit auctions: We will prove that VCG in the duplicated environment achieves a constant approximation to the expected revenue of Myerson in the original environment, and then use a single sample to "simulate" VCG with duplicates. In order to reason about the expected revenue of VCG in the duplicated environment we will utilize one of the main results of Hartline and Roughgarden's "Simple versus Optimal Mechanisms":

Lemma 7.3 ([29], Lemma 4.1)**.** *Let v_1, v_2 denote two independent and identically distributed samples from a monotone hazard rate distribution F with virtual valuation function ϕ, and t a non-negative real number. Then:*

$$\mathbb{E}\left[\max\{\phi(v_1,v_2)\}\,|\max\{(v_1,v_2)\} \geq t\right] \geq \frac{1}{3}\cdot \mathbb{E}\left[\max\{v_1,v_2\}\,|\max\{v_1,v_2\} \geq t\right]. \tag{7.13}$$

The proof is highly technical and for that reason deferred to the original paper. With this lemma, one can easily show that the expected revenue of VCG in a duplicated environment with a polymatroid constraint is a constant-factor approximation to the expected revenue of Myerson in the original environment, provided the valuation distributions satisfy the MHR condition. The proof is very similar to the one of [29, Theorem 4.2].

Theorem 7.2. *For every polymatroid environment with valuations drawn independently from distributions that satisfy the monotone hazard rate condition, the expected revenue of VCG with duplicates is at least a $\frac{1}{3}$-fraction of the expected revenue of the optimal mechanism without duplicates.*

Proof. Fix a bidder i, her duplicate i', and bids $\boldsymbol{v}_{-i}, \boldsymbol{v}'_{-i}$ for the other bidders of the duplicated environment. By the definition of the VCG mechanism, there is some threshold $t \geq 0$ such that on that valuation profile, if both v_i and $v_{i'} < t$ then both i and i' have zero allocation, and if at least one of $v_i, v_{i'}$ exceeds t, then the bidder among i, i' with higher valuation, and therefore higher virtual valuation, has non-negative allocation. Let $\boldsymbol{x}$ be the allocation of VCG. Lemma

7.3 then implies that

$$\begin{aligned}&\mathbb{E}_{v_i,v_{i'}}[\phi_i(v_i)\cdot x_i(\boldsymbol{v},\boldsymbol{v}')+\phi_i(v_{i'})\cdot x_{i'}(\boldsymbol{v},\boldsymbol{v}')\mid \boldsymbol{v}_{-i},\boldsymbol{v}'_{-i}]\\ &\geq \frac{1}{3}\cdot \mathbb{E}_{v_i,v_{i'}}[v_i\cdot x_i(\boldsymbol{v},\boldsymbol{v}')+v_{i'}\cdot x_{i'}(\boldsymbol{v},\boldsymbol{v}')\mid \boldsymbol{v}_{-i},\boldsymbol{v}'_{-i}]\end{aligned} \tag{7.14}$$

Taking expectations over $\boldsymbol{v}_{-i}, \boldsymbol{v}'_{-i}$, summing up over all pairs of duplicates, and applying linearity of expectations and Myerson's lemma yields

$$\mathbb{E}_{\boldsymbol{v},\boldsymbol{v}'}[\text{Rev}(VCG,(\boldsymbol{v},\boldsymbol{v}'))] \geq \frac{1}{3}\cdot \mathbb{E}_{\boldsymbol{v},\boldsymbol{v}'}\left[\sum_{i=1}^{2n} v_i\cdot x_i(\boldsymbol{v},\boldsymbol{v}').\right] \tag{7.15}$$

VCG always picks a social welfare-maximizing solution and the expected maximum possible social welfare in the duplicated environment, which is exactly the right-hand side of 7.15, is obviously at least that in the original environment. But the expected maximum possible social welfare in the original environment in turn upper bounds the expected revenue of any individually rational mechanism, like Myerson, in the original environment. The theorem follows. □

Once again, the auctioneer cannot feasibly "duplicate" the bidders. However, if the auctioneer is given a sample of every bidder's valuation distribution, she can use that sample to simulate VCG with duplicates.

Corollary 7.2. *For every polymatroid environment with bidders with linear valuations drawn independently from distributions that satisfy the monotone hazard rate condition, there exists a generalized single-parameter deferred-acceptance auction that, using a single sample, achieves expected revenue that is a $\frac{1}{6}$-fraction of the expected revenue of the optimal auction.*

Proof. By theorem 7.2, there exists a generalized DA auction that achieves in the duplicated environment at least one third of the expected revenue of the optimal auction in the original environment. In the duplicated environment, symmetry dictates that for every pair (i, i') of an original bidder and her duplicate, if one of them has non-negative allocation, then it is equally likely to be the original bidder or her duplicate. Thus, the expected revenue from the original bidders is at least a $\frac{1}{3}\cdot\frac{1}{2} = \frac{1}{6}$-fraction of that of the optimal auction in the original environment.
Given a sample $\boldsymbol{s} = (s_1, s_2, \ldots, s_n)$ of every bidder's valuation distribution, we can use that sample to simulate the duplication of the original environment: For every bidder i with demand d_i and valuation v_i her duplicate i' also has demand d_i and valuation s_i per unit. Then, we can run the auction described in proposition 7.1 on this extended set of bidders. □

Compared to our result for multi-unit auctions (corollary 7.1), the approximation guarantee of 7.2 is significantly worse for polymatroid environments, one sixth instead of one fourth, and it also holds for a subset of the functions that the original guarantee did, only monotone hazard rate distributions instead of all regular distributions. It is important to understand why that is the case. Multi-unit auctions, like matroid environments, are much more structured than polymatroid environments. In both of these more structured environments, there is an inherent exchange property: When comparing two maximal solutions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, whenever a bidder i had higher allocation on $\boldsymbol{x}_1$ than on $\boldsymbol{x}_2$ this implied that there was also some bidder j that had lower allocation on $\boldsymbol{x}_1$ than on $\boldsymbol{x}_2$ and that extra allocation of $\boldsymbol{x}_1$ could have instead been given to j. This coupled with the payment formula of VCG meant that whenever its allocation on some valuation profile differed from that of Myerson, the bidder with the higher allocation on VCG had to pay for her extra allocation at least the value of some bidder that was allocated more by Myerson. This exchange property, and in turn the payment property we just described, do not hold for polymatroid environments.

7.2 Multi-Parameter Environments

In this section we will propose a multi-unit deferred acceptance auction for bidders with submodular valuations that follow exclusively Gaussian and Sub-Gaussian distributions. For this environment, as discussed in chapter 6, the revenue-maximizing mechanism is not known. For this reason, as we did in that chapter, we will compare the expected revenue of the proposed mechanism directly against the expected social welfare of VCG, which is an obvious upper bound on the expected revenue of any mechanism. The setting we will study is exactly the setting with relaxed distribution assumptions of chapter 6:

The Setting (Submodular Bidders, Sub-Gaussian Distributions)

- n bidders.
- multi-unit auction with m units available for sale.
- every bidder i declares m *marginal values* $v_{i,1}, v_{i,2}, \dots, v_{i,m}$ to the mechanism, where $v_{i,j}$: Bidder i's value for acquiring her j-th unit of the good, provided she has already clinched $j-1$ units.
- every bidder's valuation function is submodular: $v_{i,1} \geq v_{i,2} \geq \dots \geq v_{i,m}$ $\forall i$
- Every marginal value $v_{i,j}$ follows some Gaussian or Sub-Gaussian distribution $F_{i,j}$.

The mechanism we propose for this setting was inspired by our own unit bundling mechanism class (definition 6.1). In our sample complexity approach we could learn the optimal bundle size and reserve price for every bidder. Now the idea is to use one sample to determine bundle sizes that are relatively close to the optimal ones and then use a second sample to learn an approximation of the bidders' valuations for those bundles. Because this mechanism makes use of two samples, we named it the two samples mechanism. Formally:

Algorithm 4: The Two Samples Mechanism ($\mathcal{TSM}$)

Collect 2 samples of all bidders' valuation distributions;
Run VCG on the first sample. Let $\boldsymbol{x} = (x_1, x_2, \dots x_n)$ be the allocation of that VCG;
For every bidder i determine her value r_i for a bundle of x_i units in the second sample;
In the auction offer to each bidder i a bundle of x_i units at a price equal to $0.85 \cdot r_i$;

Note that for each bidder i her reserve for a bundle of x_i units is not her value for that bundle in the second sample, but that number multiplied by 0.85. The reasoning behind that choice will become clear later in the analysis.
First, we will reason about the expected social welfare, on any sample, of the allocation x that was determined on the first sample.

Lemma 7.1. *In the setting described above for the expected social welfare of the allocation* $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ *that the* $\mathcal{TSM}$ *mechanism determined on the first sample, on expectation* *over the draw of the sample it holds:*

$$\mathbb{E}_{\boldsymbol{v}}\left[SW(\boldsymbol{x},\boldsymbol{v})\right] = \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i=1}^{n}\sum_{j=1}^{x_i} v_{i,j}\right] \geq \mathbb{E}_{\boldsymbol{v}}\left[SW(\mathcal{A},\boldsymbol{v})\right] - 2\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i \tag{7.16}$$

where $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_{mn}$ *is a total ordering on the standard deviations of the distributions of all* nm *marginal values.*

Remark. The notation $\mathrm{SW}(\cdot,\boldsymbol{v})$ is overloaded. If its first argument is some mechanism $\mathcal{M}$, then it represents the social welfare that that mechanism achieves on that valuation profile, according to its allocation rule, e.g. $\mathrm{SW}(\mathrm{VCG},\boldsymbol{v})$ is the social welfare of VCG on the valuation profile $\boldsymbol{v}$. If on the other hand the first argument of $\mathrm{SW}(\cdot,\boldsymbol{v})$ is some allocation $\boldsymbol{x}$, then $\mathrm{SW}(\boldsymbol{x},\boldsymbol{v})$ denotes the social welfare of that allocation on that valuation profile.

Proof. The proof is an extension of theorem 6.5. The distribution assumptions are the same as they were in that theorem. There, we proved that on expectation, the m out of the nm marginal values that will "overshoot" (i.e. be higher than expected) the most their expected values on the first sample $\boldsymbol{v}^1$ will in total be at most $\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i$ higher than their expected value. With an identical argument, one can prove that on expectation, the m out of the nm marginal values that will "undershoot" (i.e. be lower than expected) the most their expected values on that same sample $\boldsymbol{v}^1$ will in total be at most $\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i$ lower than their total expected value. These statements also hold for the allocation $\boldsymbol{x}$ that $\mathcal{TSM}$ determined on the sample and the fixed allocation of $\mathcal{A}$:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\boldsymbol{x},\boldsymbol{v})] &\geq \mathrm{SW}(\boldsymbol{x},\boldsymbol{v}^1) - \sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i \\
&\geq \mathrm{SW}(\mathcal{A},\boldsymbol{v}^1) - \sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i \\
&\geq \left(\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathcal{A},\boldsymbol{v})] - \sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i\right) - \sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i \\
&= \mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathcal{A},\boldsymbol{v})] - 2\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i
\end{aligned} \tag{7.17}$$

where the first inequality follows from the fact that, as we explained, the social welfare of allocation $\boldsymbol{x}$ will on expectation be at most $\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i$ higher on the sample $\boldsymbol{v}^1$ than its expected value and the second inequality follows from the fact that the allocation $\boldsymbol{x}$ is the result of running VCG on $\boldsymbol{v}^1$, so on that sample it has the highest possible social welfare of any allocation, including the allocation of $\mathcal{A}$. Finally, the third inequality follows from the fact that the allocation of $\mathcal{A}$, just like any other allocation, on expectation over the draw of the sample $\boldsymbol{v}^1$ can be at most $\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i$ lower on $\boldsymbol{v}^1$ than its expected value. □

Lemma 7.1 shows that the expected social welfare of the allocation $\boldsymbol{x}$ that $\mathcal{TSM}$ determined on the first sample is close to the expected social welfare of $\mathcal{A}$, which in turn is close to the expected social welfare of VCG. The next step is to show how close the expected revenue of $\mathcal{TSM}$ is to the expected social welfare of $\boldsymbol{x}$. The following 2 lemmata do exactly that.
In order to prove our next lemma we will need one of the main results from Roughgarden's and Huang's "Making the Most of Your Samples".

Theorem 7.3 ([31], Theorem 5.1)**.** *In the case of a single item available for sale and a single buyer, setting a reserve price equal to* 0.85 *times the bidder's valuation for the item on a single sample is* 0.589*-approximate to the optimal expected revenue for monotone hazard rate distributions.*

Providing an intuitive explanation for this result is not possible, as the proof is quite involved. For this reason, we will give an intuitive explanation for another, more basic result that is however in the same vein.

Theorem 7.4. *In the case of a single item available for sale and a single buyer, setting a reserve price equal to the bidder's valuation for the item on a single sample is* $\frac{1}{2}$*-approximate to the optimal expected revenue for regular distributions.*

Proof Sketch. Let D be the regular distribution of the valuation of the single bidder, and $R(q)$ be its revenue curve in probability space: For any $q \in [0,1]$ the value of $R(q)$ is the expected revenue of posting a price such that the probability that the bidder with valuation distribution D will purchase the item is q. If $v(q)$ is that price, then the revenue curve in probability space is simply $R(q) = qv(q)$. For regular distributions. The optimal reserve price r^*, also corresponds to the optimal probability $q^* = v(r)$ that maximizes the expected revenue, and consequently, $R(q)$. Pictorially, the value of the expected revenue at the optimal reserve price corresponds exactly to the height at $R(q^*)$, which is the peak of the revenue curve. By setting a reserve price equal to a random sample from the distribution, the expected revenue at that point will be equal to a random point of $R(q)$, with all points being equally likely, i.e. the expected revenue will be $\int_0^1 R(q)dq$. Pictorially, this is the total area under the curve $R(q)$. But for regular distributions, the revenue curve in probability space is convex. This means that $R(q)$ is at least a triangle, with its highest point being $R(q)$. But the area of a triangle with height $R(q^*)$ and width 1 (the range of values that q can take) is exactly $R(q^*) \cdot \frac{1}{2} = \frac{R(q^*)}{2}$, i.e. half of the expected revenue at the optimal reserve price. □

Lemma 7.4. *Let* $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ *be the allocation that the Two Samples Mechanism* ($\mathcal{TSM}$) *determined on the first sample. Then, if all the bidders' distributions satisfy the monotone hazard condition, the expected revenue of* $\mathcal{TSM}$ *is at least a* 0.589*-fraction of the expected revenue of posting to each bidder* i *the optimal price* r_i^* *for a bundle of* x_i *units.*

Proof. Fix an arbitrary bidder i. Let $\widetilde{F}_i$ be the monotone hazard rate distribution of bidder i's valuation for receiving x_i units and r_i^* the optimal reserve price for that bundle of x_i units. Let $\widehat{R}_i(p) = p(1 - \widetilde{F}_i(p))$ denote bidder i's revenue function for x_i units in value space. Finally, for notational simplicity let $v_i^1(x_i)$ be bidder i's value for x_i units in the first sample, $\boldsymbol{v}^1$, i.e. $v_i^1(x_i) = \sum_{j=1}^{x_i} v_{i,j}^1$. The key observation is that the price r_i that the Two Samples Mechanism will set for that bundle of units offered to bidder i is determined by the second sample, thus it is a random, independent sample from the distribution $\widetilde{F}_i$. Applying theorem 7.3, we have:

$$\mathbb{E}_{v_i^1(x_i)}\left[\widehat{R}_i(0.85 \cdot v_i^1(x_i))\right] \geq 0.589 \cdot \widehat{R}_i(r_i^*) \tag{7.18}$$

Summing up over all bidders and using linearity of expectations:

$$\sum_{i=1}^{n} \mathbb{E}_{v_i^1(x_i)}\left[\widehat{R}_i(0.85 \cdot v_i^1(x_i))\right] \geq 0.589 \cdot \sum_{i=1}^{n} \widehat{R}_i(r_i^*) \implies$$

$$\mathbb{E}_{\boldsymbol{v}^1}\left[\sum_{i=1}^{n} \widehat{R}_i(0.85 \cdot v_i(x_i))\right] \geq 0.589 \cdot \sum_{i=1}^{n} \widehat{R}_i(r_i^*) \tag{7.19}$$

□

Lemma 7.4 proves that the expected revenue of the two samples mechanism is a constant approximation of the expected revenue if for those bundles, the auctioneer had set the optimal reserve prices. The final step is to show that for those optimal reserve prices, the expected revenue is close to the expected social welfare of the allocation $\boldsymbol{x}$ determined on the first sample.

Lemma 7.5. *Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ be the allocation determined by the Two Samples Mechanism on the first sample. Let $\widetilde{B}_i, \widetilde{\sigma_i}^2$ be the expected value and variance of $\widetilde{F}_i$, bidder i's valuation distribution for x_i units. Then, if all the valuation distributions are monotone hazard rate, on expectation over the draw of the two samples $\boldsymbol{v}^1, \boldsymbol{v}^2$ that $\mathcal{TSM}$ used the following bound holds:*

$$\mathbb{E}_{\boldsymbol{v}}[Rev(\mathcal{TSM}, \boldsymbol{v})] \geq 0.589 \cdot \mathbb{E}_{\boldsymbol{v}}[SW(\boldsymbol{x}, \boldsymbol{v})] - 0.589 \cdot \sum_{i=1}^{n} (3\widetilde{B}_i)^{1/3} \widetilde{\sigma_i}^{2/3} \tag{7.20}$$

Proof. The proof is an application of lemma 7.4, corollary 6.1 and linearity of expectations:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{v}}[\mathrm{Rev}(\mathcal{TSM}, \boldsymbol{v})] &= \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i=1}^{n} \widehat{R}_i(0.85 \cdot v_i(x_i))\right] \\
&\geq 0.589 \cdot \sum_{i=1}^{n} \widehat{R}(r_i^*) \\
&= 0.589 \cdot \sum_{i=1}^{n} \left[\mathbb{E}_{v_i(x_i) \sim \widetilde{F}_i}[v_i(x_i)] - S_i\right] \\
&= 0.589 \cdot \sum_{i=1}^{n} \mathbb{E}_{v_i(x_i) \sim \widetilde{F}_i}[v_i(x_i)] - 0.589 \cdot \sum_{i=1}^{n} S_i \\
&\geq 0.589 \cdot \sum_{i=1}^{n} \widetilde{B}_i - 0.589 \cdot \sum_{i=1}^{n} (3\widetilde{B}_i)^{1/3} \widetilde{\sigma_i}^{2/3} \\
&= 0.589 \cdot \mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\boldsymbol{x}, \boldsymbol{v})] - 0.589 \cdot \sum_{i=1}^{n} (3\widetilde{B}_i)^{1/3} \widetilde{\sigma_i}^{2/3}
\end{aligned} \tag{7.21}$$

where the first inequality follows from lemma 7.4, the second equality follows from the definition of separation and the third equality from linearity of expectations. The final inequality follows from corollary 6.1 and the final equality from the fact that the expected social welfare of the allocation $\boldsymbol{x}$ is simply the sum of the expected value of every bidder for the units that she is allocated in $\boldsymbol{x}$.

□

Combining lemmata 7.1 and 7.5 with the lower bound on the expected social welfare of $\mathcal{A}$ of theorem 6.5 yields the following corollary:

Corollary 7.3. *In the setting of multi-unit auctions with submodular bidders whose valuations follow Gaussian and Sub-Gaussian distributions, as described at the start of the section, for the expected revenue of the Two Samples Mechanism, on expectation over the draw of the two samples that determined the bundle sizes and reserve prices it holds:*

$$\mathbb{E}_{\boldsymbol{v}}[Rev(\mathcal{TSM}, \boldsymbol{v})] \geq 0.589 \cdot \mathbb{E}_{\boldsymbol{v}}[SW(VCG, \boldsymbol{v})] - 1.767\sqrt{2\log(nm)} \sum_{i=1}^{m} \overline{\sigma}_i - 0.589 \cdot \sum_{i=1}^{n} (3\widetilde{B}_i)^{1/3} \widetilde{\sigma_i}^{2/3} \tag{7.22}$$

where $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_{mn}$ is a total ordering on the standard deviations of the distributions of all nm marginal values and $\widetilde{B}_i, \widetilde{\sigma_i}^2$ are the expected value and variance of $\widetilde{F}_i$, bidder i's valuation distribution for x_i units, i.e. $\widetilde{\sigma_i}^2 = \sum_{j=1}^{x_i} \sigma_{i,j}^2$.

Proof. For the expected revenue of the Two Samples Mechanism, on expectation over the draw of the two samples it used, it holds:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{v}}[\text{Rev}(\mathcal{TSM}, \boldsymbol{v})] &\geq 0.589 \cdot \mathbb{E}_{\boldsymbol{v}}[\text{SW}(\boldsymbol{x}, \boldsymbol{v})] - 0.589 \cdot \sum_{i=1}^{n} (3\widetilde{B}_i)^{1/3} \widetilde{\sigma_i}^{2/3} \\
&\geq 0.589 \cdot \left(\mathbb{E}_{\boldsymbol{v}}[\text{SW}(\mathcal{A}, \boldsymbol{v})] - 2\sqrt{2\log(nm)} \sum_{i=1}^{m} \overline{\sigma}_i \right) - 0.589 \cdot \sum_{i=1}^{n} (3\widetilde{B}_i)^{1/3} \widetilde{\sigma_i}^{2/3} \\
&\geq 0.589 \cdot \left(\mathbb{E}_{\boldsymbol{v}}[\text{SW}(\text{VCG}, \boldsymbol{v})] - 3\sqrt{2\log(nm)} \sum_{i=1}^{m} \overline{\sigma}_i \right) - 0.589 \cdot \sum_{i=1}^{n} (3\widetilde{B}_i)^{1/3} \widetilde{\sigma_i}^{2/3} \\
&= 0.589 \cdot \mathbb{E}_{\boldsymbol{v}}[\text{SW}(\text{VCG}, \boldsymbol{v})] - 1.767\sqrt{2\log(nm)} \sum_{i=1}^{m} \overline{\sigma}_i - 0.589 \cdot \sum_{i=1}^{n} (3\widetilde{B}_i)^{1/3} \widetilde{\sigma_i}^{2/3}
\end{aligned} \tag{7.23}$$

where the first inequality follows from lemma 7.5, the second from lemma 7.1 and the final one from theorem 6.5.

□

It is noteworthy that under the exact same distribution assumptions that we had used in our sample complexity approach, we managed to propose a mechanism that, given only 2 samples of the bidders' valuation distributions, achieves expected revenue that is a 0.589-approximation of the expected revenue that the unit-bundling class achieves using a logarithmic to the number of bidders number of samples, minus an additional $1.178\sqrt{2\log(nm)} \sum_{i=1}^{m} \overline{\sigma}_i$. For high-valued items, this last term becomes trivially small compared to the expected revenue of the auction. There is an interesting final remark to be made. Both in the case of budget-additive bidders and in the case of submodular ones, when restricting the number of samples we are still picking mechanisms from the corresponding mechanism class that we had proposed for the same environment in our sample complexity approach: In the case of budget-additive bidders, we use a single sample to determine an approximation of the threshold of a 1-level auction for every bidder, while in the case of submodular bidders we are using one sample to determine the bundle sizes and another to determine the reserve prices of a mechanism from the $\mathcal{UB}$ class (definition 6.1).

7.2.1 Generalizing the Previous Result for More Samples

It is possible that the auctioneer, in this setting, managed to procure a constant number of samples that is higher than 2 yet not high enough so that the empirical revenue maximizer from the $\mathcal{UB}$ class generalizes well. A question that naturally arises in this case is whether the auctioneer could somehow utilize those extra samples, or she is restricting in utilizing only two of them and fielding the two samples mechanism. Fortunately, the Two Samples Mechanism can easily be extended to allow for more samples. We named the resulting mechanism the "Few Samples Mechanism". Formally:

Algorithm 5: The Few Samples Mechanism ($\mathcal{TSM}$)

1 Collect $N \geq 2$ samples of all bidders' valuation distributions;
2 Randomly exclude 1 sample. Use the $N-1$ samples left to calculate the average value of all marginal values ;
3 Run VCG on the average marginal values calculated in step 2. Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ be the allocation of that VCG ;
4 For every bidder i determine her value r_i for a bundle of x_i units in the sample that was excluded in step 2;
5 In the auction offer to each bidder i a bundle of x_i units at a price equal to $0.85 \cdot r_i$;

Essentially, this mechanism now uses $N-1$ samples to determine the bundle sizes that will offer to each bidder, while $\mathcal{TSM}$ used only one. Intuitively, this more informed decision results in bundle sizes that more closely resemble the allocation of $\mathcal{A}$. The following lemma formalizes this intuition.

Lemma 7.6. *In the setting described at the start of this section, for the expected social welfare of the allocation $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ that the Few Samples Mechanism determined on $N-1$ out of the N samples, on expectation over the draw of those samples it holds:*

$$\mathbb{E}_{\boldsymbol{v}}\left[SW(\boldsymbol{x}, \boldsymbol{v})\right] = \mathbb{E}_{\boldsymbol{v}}\left[\sum_{i=1}^{n}\sum_{j=1}^{x_i} v_{i,j}\right] \geq \mathbb{E}_{\boldsymbol{v}}\left[SW(\mathcal{A}, \boldsymbol{v})\right] - 2\sqrt{\frac{2\log(nm)}{N-1}}\sum_{i=1}^{m}\overline{\sigma}_i \tag{7.24}$$

where $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_{mn}$ is a total ordering on the standard deviations of the distributions of all nm marginal values.

The proof is very similar to that of lemma 7.1. The main difference is that now that the average values of more samples are used to determine the bundle sizes offered to the bidders, the variances of the distributions of those averages on the samples, if viewed as random variables, are significantly lower than the variances of the original distributions.

Proof. In this setting, all marginal values follow exclusively Gaussian and Sub-Gaussian distributions with expected values $\mu_1, \mu_2, \ldots, \mu_{mn}$ and standard deviations/Sub-Gaussian parameters $\sigma_1, \sigma_2, , \ldots, \sigma_{mn}$ respectively. Then, their respective averages out of $N-1$ samples follow Gaussian and Sub-Gaussian distributions with expected values $\mu_1, \mu_2, \ldots, \mu_{mn}$ and variances/ Sub-Gaussian parameters at most $\frac{\sigma_1^2}{N-1}, \frac{\sigma_2^2}{N-1}, \ldots, \frac{\sigma_{mn}^2}{N-1}$. This implies that their standard deviations/Sub-Gaussian parameters are at most $\frac{\sigma_1}{\sqrt{N-1}}, \frac{\sigma_2}{\sqrt{N-1}}, , \ldots, \frac{\sigma_{mn}}{\sqrt{N-1}}$. With an argument identical to that of lemma 7.1, now the most that the averages of m marginal values can either "overshoot" or "undershoot" (i.e. be higher or lower than expected) their expected values is

$$\sqrt{2\log(nm)}\sum_{i=1}^{m}\frac{\overline{\sigma}_i}{\sqrt{N-1}} = \sqrt{\frac{2\log(nm)}{N-1}}\sum_{i=1}^{m}\overline{\sigma}_i. \tag{7.25}$$

Once again, these statements also hold for the allocation $\boldsymbol{x}$ that $\mathcal{FSM}$ determined on the sample and the fixed allocation of $\mathcal{A}$. Let $\hat{\boldsymbol{v}}$ be the vector of the empirically average marginal values, as determined by $\mathcal{FSM}$ on the $N-1$ samples. Then

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\boldsymbol{x},\boldsymbol{v})] &\geq \mathrm{SW}(\boldsymbol{x},\hat{\boldsymbol{v}}) - \sqrt{\frac{2\log(nm)}{N-1}}\sum_{i=1}^{m}\overline{\sigma}_i \\
&\geq \mathrm{SW}(\mathcal{A},\hat{\boldsymbol{v}}) - \sqrt{\frac{2\log(nm)}{N-1}}\sum_{i=1}^{m}\overline{\sigma}_i \\
&\geq \left(\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathcal{A},\boldsymbol{v})] - \sqrt{\frac{2\log(nm)}{N-1}}\sum_{i=1}^{m}\overline{\sigma}_i\right) - \sqrt{\frac{2\log(nm)}{N-1}}\sum_{i=1}^{m}\overline{\sigma}_i \\
&= \mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathcal{A},\boldsymbol{v})] - 2\sqrt{\frac{2\log(nm)}{N-1}}\sum_{i=1}^{m}\overline{\sigma}_i
\end{aligned} \tag{7.26}$$

where once again the first inequality follows from the fact that, as we explained, the social welfare of allocation $\boldsymbol{x}$ will on expectation be at most $\sqrt{2\log(nm)/(N-1)}\sum_{i=1}^{m}\overline{\sigma}_i$ higher on $\hat{\boldsymbol{v}}$ than its expected value and the second inequality follows from the fact that the allocation $\boldsymbol{x}$ is the result of running VCG on $\hat{\boldsymbol{v}}$, so on that valuation profile it has the highest possible social welfare of any allocation, including the allocation of $\mathcal{A}$. Finally, the third inequality follows from the fact that the allocation of $\mathcal{A}$, just like any other allocation, on expectation over the valuation profile $\hat{\boldsymbol{v}}$ can be at most $\sqrt{2\log(nm)/(N-1)}\sum_{i=1}^{m}\overline{\sigma}_i$ lower on $\hat{\boldsymbol{v}}$ than its expected value. □

As discussed earlier, lemma 7.6 yields a better approximation guarantee for the expected social welfare of the allocation $\boldsymbol{x}$ than lemma 7.1 in the case where more samples are available. Combining it (instead of lemma 7.1) with lemma 7.5 and the lower bound on the expected social welfare of $\mathcal{A}$ of theorem 6.5 yields the following corollary:

Corollary 7.4. *In the setting of multi-unit auctions with submodular bidders whose valuations follow (MHR) Gaussian/Sub-Gaussian distributions, as described at the start of the section, for the expected revenue of the modified Two Samples Mechanism, on expectation over the draw of the two samples that determined the bundle sizes and reserve prices it holds:*

$$\mathbb{E}_{\boldsymbol{v}}[Rev(\mathcal{TSM},\boldsymbol{v})] \geq 0.589\cdot\left(\mathbb{E}_{\boldsymbol{v}}[SW(VCG,\boldsymbol{v})] - \left(\frac{2}{N-1}+1\right)\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i - \sum_{i=1}^{n}(3\widetilde{B}_i)^{1/3}\widetilde{\sigma}_i^{2/3}\right) \tag{7.27}$$

where $\overline{\sigma}_1 \geq \overline{\sigma}_2 \geq \cdots \geq \overline{\sigma}_{mn}$ is a total ordering on the standard deviations of the distributions of all nm marginal values and $\widetilde{B}_i, \widetilde{\sigma}_i^2$ are the expected value and variance of $\widetilde{F}_i$, bidder i's valuation distribution for x_i units, i.e. $\widetilde{\sigma}_i^2 = \sum_{j=1}^{x_i}\sigma_{i,j}^2$.

Proof. For the expected revenue of the Few Samples Mechanism, on expectation over the draw of the N samples it used, it holds:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{v}}[\mathrm{Rev}(\mathcal{FSM},\boldsymbol{v})] &\geq 0.589\cdot\left(\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\boldsymbol{x},\boldsymbol{v})] - \sum_{i=1}^{n}(3\widetilde{B}_i)^{1/3}\widetilde{\sigma}_i^{2/3}\right) \\
&\geq 0.589\cdot\left(\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathcal{A},\boldsymbol{v})] - 2\sqrt{\frac{2\log(nm)}{N-1}}\sum_{i=1}^{m}\overline{\sigma}_i - \sum_{i=1}^{n}(3\widetilde{B}_i)^{1/3}\widetilde{\sigma}_i^{2/3}\right) \\
&\geq 0.589\cdot\left(\mathbb{E}_{\boldsymbol{v}}[\mathrm{SW}(\mathrm{VCG},\boldsymbol{v})] - \left(\frac{2}{N-1}+1\right)\sqrt{2\log(nm)}\sum_{i=1}^{m}\overline{\sigma}_i - \sum_{i=1}^{n}(3\widetilde{B}_i)^{1/3}\widetilde{\sigma}_i^{2/3}\right)
\end{aligned} \tag{7.28}$$

where the first inequality follows from lemma 7.5, the second from lemma 7.1 and the final one from theorem 6.5.

□

At a first glance, this result may be hard to grasp. The few samples mechanism, using some small constant N number of samples achieves expected revenue that is a 0.589-approximation of the expected revenue that the unit-bundling class achieves using a logarithmic to the number of bidders number of samples, minus an additional additive term, $1.178\sqrt{2\log(nm)/(N-1)}\sum_{i=1}^{m}\overline{\sigma}_i$. As discussed earlier, for high-valued goods, this term becomes trivially small compared to the expected revenue of the mechanism.

Chapter 8

Conclusion and Future Work

In this work we explored the revenue performance of deferred-acceptance auctions for different environments. Specifically, in the case of multi-unit auctions with budget-additive bidders, as well as more general environments with polymatroid constraints, under natural distribution assumptions, we proposed a novel DA auction mechanism class and upper-bounded the sample-complexity of learning a revenue-optimal auction from that class. Then, in the case of bidders with submodular valuations, we proposed another DA auction class and lower-bounded the expected revenue of the empirical revenue maximizer of that class, in relation to the number of samples used. Finally, we explored how both of these results change in the case where the auctioneer only possesses a very restricted number of samples. The most meaningful question that arises from our work, apart of course of improving the bounds we proved, is whether, and how, these results generalize in the case where there are different goods available for sale. We conjure that for bidders with additive and submodular valuation functions, similar results can be proven, albeit with a higher sample complexity due to the more complex structure of the valuation functions.

Bibliography

[1] Lawrence Ausubel and Paul Milgrom. 2004. The Lovely but Lonely Vickrey Auction. In Discussion Papers 03-036, Stanford Institute for Economic Policy Research.

[2] Noam Nisan et al. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[3] Jason D. Hartline. *Mechanism Design and Approximation*.

[4] Tim Roughgarden. *Twenty Lectures on Algorithmic Game Theory*. Cambridge University Press, 2016.

[5] David Pollard. *Convergence of Stochastic Processes*. Springer, 1984.

[6] Roger Myerson. 1981. Optimal auction design. In Mathematics of Operations Research, 6(1):58-73.

[7] Edward H. Clarke. 1971. Multipart pricing of public goods. In Public Choice, 11(1):17-33.

[8] Theodore Groves. 1973. Incentives in teams. Econometrica, 41(4):617-631.

[9] William Vickrey. 1961. Counterspeculation, auctions, and competitive sealed tenders. In Journals of Finance, 16(1):8-37.

[10] Tim Roughgarden et al. 2013. CS364A: Algorithmic Game Theory Lecture #5: Revenue-Maximizing Auctions.

[11] Tim Roughgarden et al. 2013. CS364A: Algorithmic Game Theory Lecture #7: Multi-Parameter Mechanism Design and the VCG Mechanism.

[12] Maria-Florina Balcan et al. 2018. A General Theory of Sample Complexity for Multi-Item Profit Maximization. In ACM EC 2018 - Proceedings of the 2018 ACM conference on Economics and Computation.

[13] Jamie Morgenstern and Tim Roughgarden. 2016. The Pseudo-Dimension of Near-Optimal Auctions. In Advances in Neural Processing Systems.

[14] Paul Milgrom and Ilya Segal. 2014. Deferred-acceptance auctions and radio spectrum reallocation. In ACM Conference on Economics and Computation, EC '14, Stanford , CA, USA.

[15] Paul Milgrom and Ilya Segal. 2013. Deferred-Acceptance Heuristic Auctions.

[16] Shengwu Li. 2017. Obviously Strategy-Proof Mechanisms. In American Economic Review, 107 (11): 3257-87.

[17] FCC. 2017. Broadcast Incentive Auction. In https://www.fcc.gov/about-fcc/fcc-initiatives/incentive-auctions.

[18] Vasilis Gkatzelis, Evangelos Markakis and Tim Roughgarden. 2017. In Proceedings of the 2017 ACM Conference on Economics and Computation, 1(1): 21-38.

[19] Bernhard Korte and Jens Vygen, 2007. Combinatorial Optimazation: Theory and Algorithms. Springer Verlag, Berlin Heidelberg, Germany.

[20] Paul Dutting, Vasilis Gkatzelis and Tim Roughgarden. 2014. The Performance of Deferred-Acceptance Auctions. In ACM Conference on Economics and Computation, EC'14, 187-204.

[21] Nikhil R. Devanur, Milena Mihail and Vijay Vazirani. 2005. Strategyproof Cost-Sharing Mechanisms for Set Cover and Facility Location Games. In Decision Support Systems 39, 11-22.

[22] Tim Roughgarden and Mukund Sundararajan. 2009. Quantifying Inefficiency in Cost-Sharing Mechanisms. Journal of the ACM 56, 4.

[23] Jamie Morgenstern and Tim Roughgarden. 2015. In Advances in Neural Information Processing Systems 2015, 136-144.

[24] Sushil Bikhchandani et al. 2011. An Ascending Vickrey Auction for Selling Bases of a Matroid. In Operations Research 59 (2011), 400-413.

[25] Andres Medina and Sergei Vassilvitskii. Revenue Optimization with Approximate Bid Predictions. In 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

[26] Pascal Massart. Concentration Inequalities and Model Selection. Springer, 1971.

[27] George Casella and Roger Lee Berger. *Statistical Inference*. Cengage Learning, 2008.

[28] Aidan Lyon. 2013. Why are Normal Distributions Normal? In The British Journal for the Philosophy of Science, Volume 65, Issue 3, September 2014, Pages 621–649.

[29] Jason D. Hartline and Tim Roughgarden. 2009. Simple versus Optimal Mechanisms. In EC '09: Proceedings of the 10th ACM conference on Electronic commerce, pages 225-234.

[30] Satoru Fujishige. 1980. Lexicographically Optimal Base of a Polymatroid with Respect to a Weight Vector. In Mathematics of Operations Research Vol. 5, No. 2 (May, 1980), pages 186-196.

[31] Zhiyi Huang, Yishay Mansour and Tim Roughgarden. 2018. Making the Most of Your Samples. In SIAM Journal on Computing 47(3):651-674.

[32] Vincent Conitzer and Tuomas Sandholm. 2002. Complexity of mechanism design. In Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI'02), pages 103-110.

[33] Vincent Conitzer and Tuomas Sandholm. 2003. Applications of automated mechanism design. In Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI).

[34] Vincent Conitzer and Tuomas Sandholm. 2004. Self-interested automated mechanism design and implications for optimal combinatorial auctions. In Proceedings of the ACM Conference on Economics and Computation (EC), pages 132–141.

[35] Anton Likhodedov and Tuomas Sandholm. 2004. Methods for boosting revenue in combinatorial auctions. In Proceedings of the AAAI Conference on Artificial Intelligence.

[36] Anton Likhodedov and Tuomas Sandholm. 2005. Approximating revenue-maximizing combinatorial auctions. In Proceedings of the AAAI Conference on Artificial Intelligence.

[37] Tuomas Sandholm and Anton Likhodedov. 2015. Automated design of revenue-maximizing combinatorial auctions. In Operations Research, 63(5): pages 1000–1025.

[38] Maria-Florina Balcan, Tuomas Sandholm and Ellen Vitercik. 2016. Sample complexity of automated mechanism design. In Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS).

[39] Yang Cai and Constantinos Daskalakis. 2017. Learning Multi-item Auctions with (or without) Samples.In Proceedings of the IEEE Symposium on Foundations of Computer Science (FOCS).