



**HAL**  
open science

# Bidding efficiently in Simultaneous Ascending Auctions using Monte Carlo Tree Search

Alexandre Pacaud

► **To cite this version:**

Alexandre Pacaud. Bidding efficiently in Simultaneous Ascending Auctions using Monte Carlo Tree Search. Computer Science and Game Theory [cs.GT]. Institut Polytechnique de Paris, 2024. English. NNT : 2024IPPAT003 . tel-04520259

**HAL Id: tel-04520259**

**<https://theses.hal.science/tel-04520259v1>**

Submitted on 25 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Bidding efficiently in Simultaneous Ascending Auctions using Monte Carlo Tree Search

Thèse de doctorat de l'Institut Polytechnique de Paris  
préparée à Télécom ParisTech

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (ED  
IP PARIS)  
Spécialité de doctorat: Mathématiques et informatique

Thèse présentée et soutenue à Châtillon, le 19 janvier 2024, par

**ALEXANDRE PACAUD**

Composition du Jury :

Simon Lucas Professeur, Queen Mary University of London	Président/Examineur
Tristan Cazenave Professeur, LAMSADE, Université Paris-Dauphine	Rapporteur
Patrick Maillé Professeur, IMT Atlantique	Rapporteur
Emilie Kaufmann CNRS Researcher, CRISAL, CNRS	Examineur
Marceau Coupechoux Professeur, INFRES, Telecom ParisTech	Directeur de thèse
Aurélien Bechler Ingénieur de Recherche, Orange Labs	Encadrant



# Acknowledgements

I am deeply grateful to my supervisors, Aurélien and Marceau, for their assistance during my three-year PhD journey. Their guidance has offered valuable insights, and I hope they have found our collaboration equally rewarding.

I extend my sincere thanks to all the jury members, particularly the reviewers, for dedicating their time to evaluate my manuscript and travelling to Paris for my PhD defence. I understand that it was quite a distance for some of you.

I want to express my gratitude to Mikael for being a real support throughout my thesis and for giving me the opportunity to assist him with the Game Theory course at CentraleSupélec. I also extend my gratitude to my colleagues at Orange, including Olfa, Pierre, Matthieu, Taoufik, Stéphane, Guillaume, Joshua, Morgan, and many others, who have contributed significantly to making my PhD experience enjoyable.

Special thanks to Marine and Nawel, as well as the entire Communication team at Orange, for their efforts in organising "Ma Thèse en trois minutes" and the Orange learning expedition to San Francisco. I would also like to thank all the participants, and particularly Remy and Nicolas, with whom I shared great laughs.

Finally, I would like to express my gratitude to my friends and family, who make my everyday life fantastic. I want to give a special thanks to my parents and my sister for making my life so easy and for supporting me over the past 27 years. I also want to thank my flatmate, Melvin, for all the great moments we have shared during these three years.

To all my dear friends who attended my defence, I am grateful for your presence, listed in the order of your arrival: Julia, Alix, Hugo, Ilyass, Benoît, Pierre, Jean, Agathe, Geoffroy, and Nicholas. Additionally, I want to acknowledge my other friends who could not attend but remain important to me: Sebastian, Nathan, Bertrand, Thomas, Matthew, Meghan, Gaëlle, Francesca, Juliette, Alexandre, Ulysse, Marine, Eloïse, Thanh-ai, Tristan, Thibault, Guillaume, Paul, and many more.



# Résumé détaillé en français

## Introduction

Au fil des dernières décennies, les opérateurs mobiles ont mené des batailles acharnées pour obtenir du spectre cellulaire, cherchant ainsi à améliorer leur qualité de service et à développer leurs réseaux de communication sans fil. À l'image d'un peintre qui assemble différentes couleurs pour créer ses plus belles œuvres, un opérateur mobile a besoin d'une vaste palette de fréquences pour offrir la meilleure qualité de service possible. Par exemple, les fréquences élevées offrent des débits plus élevés que les fréquences basses. Cependant, elles transmettent les informations sur des distances plus courtes et ne traversent pas aussi efficacement les obstacles denses. Ainsi, pour rester compétitif et garantir une qualité de service optimale, il est crucial pour un opérateur d'avoir accès aux deux bandes de fréquences.

De nos jours, la plupart des pays attribuent le spectre cellulaire sous forme de licences via des enchères. Au cours des vingt dernières années, l'enchère ascendante simultanée (SAA) s'est imposée comme le mécanisme privilégié pour la vente du spectre cellulaire licencié. Cette enchère a la particularité d'avoir un format dynamique en plusieurs tours, où les joueurs soumettent simultanément leurs offres sur toutes les licences. L'enchère se termine lorsqu'aucune nouvelle offre a été soumise durant un tour. Sa popularité repose sur ses règles relativement simples ainsi que sur la génération de revenus substantiels pour le régulateur. Elle a notamment été récemment utilisée en Allemagne, au Portugal et en Italie pour l'attribution des fréquences 5G. De plus, il est très probable que le SAA joue un rôle central dans l'attribution des fréquences 6G prévue aux alentours de 2030.

Au vu des sommes d'argent mises en jeu, parfois dépassant le milliard d'euros, et des implications stratégiques déterminées par les résultats de cette enchère, il est fondamental

pour les opérateurs mobiles d'avoir une bonne stratégie d'enchérissement. Cependant, en raison de la complexité intrinsèque élevée du jeu associé au SAA, la théorie des enchères et les méthodes exactes de résolution de jeux ne peuvent pas calculer la stratégie d'enchérissement optimale. De plus, le SAA engendre divers problèmes stratégiques en raison de son mécanisme particulier et des caractéristiques spécifiques des opérateurs, tels que les contraintes budgétaires, ajoutant ainsi une couche de complexité supplémentaire à son étude.

Dans cette thèse, nous proposons une solution pour enchérir efficacement dans le SAA fondée sur la recherche arborescente de Monte Carlo (MCTS) ainsi que sur une prédiction spécifique des prix finaux. Notre algorithme traite simultanément les principaux problèmes stratégiques du SAA. Afin de concevoir un tel algorithme, nous avons graduellement augmenté la complexité de notre modèle SAA au fil des chapitres. La thèse est structurée en six chapitres.

Le chapitre 1 présente les différents mécanismes d'enchère utilisés au fil des ans pour la vente du spectre cellulaire licencié, mettant en évidence leurs avantages et leurs inconvénients. Dans le chapitre 2, nous approfondissons le problème de l'enchérisseur au sein du SAA, examinons les travaux pertinents visant à résoudre ce problème, et présentons nos propres contributions. Le chapitre 3 passe en revue les différentes méthodes traditionnelles de recherche dans les jeux avec des adversaires, en mettant particulièrement l'accent sur le MCTS. Le chapitre 4 considère un modèle déterministe du SAA au tour par tour à information parfaite et complète. Nous introduisons un algorithme MCTS performant nommé  $MS^\lambda$  pour enchérir dans ce jeu d'enchères ainsi qu'une nouvelle méthode pour prédire les prix finaux des objets. Dans le chapitre 5, nous étendons le modèle précédent à la soumission simultanée des offres, aux contraintes budgétaires et aux règles d'activité, tout en maintenant le jeu à information complète. Nous adaptons notre méthode de prédiction des prix finaux du chapitre précédent et présentons un nouvel algorithme MCTS performant nommé  $SM S^\alpha$ . Les expériences numériques sont menées sur des instances de taille réelle. Dans le chapitre 6, nous considérons le jeu précédent sous sa version à information incomplète pour modéliser les différentes incertitudes présentes dans la réalité. Nous présentons alors trois algorithmes performants qui adaptent  $SM S^\alpha$  à ce nouveau problème.

## Une introduction aux enchères du spectre licencié

L'attribution du spectre cellulaire aux opérateurs se déroule en deux étapes. Dans un premier temps, le spectre est associé à des licences, chaque licence étant définie par sa bande de fréquence, sa couverture géographique, sa période d'utilisation et ses restrictions d'utilisation. Ensuite, lors de la deuxième étape, les licences sont attribuées aux différentes entreprises. Trois méthodes sont couramment utilisées à cette fin : les concours de beauté, où chaque participant soumet une proposition et le régulateur choisit de manière arbitraire l'allocation des licences, les loteries et les enchères. Actuellement, les licences sont principalement attribuées par le biais d'enchères, car cette procédure est transparente et permet généralement d'attribuer les licences aux opérateurs les mieux adaptés pour développer le réseau. Une enchère a principalement deux objectifs : attribuer les licences aux opérateurs les mieux adaptés pour développer le réseau et maximiser les revenus du régulateur. Pour atteindre ces deux objectifs, le régulateur doit choisir judicieusement son mécanisme d'enchères.

Plusieurs mécanismes ont été utilisés au fil des années pour la vente de licences. On peut les diviser en deux groupes : les enchères à un tour et les enchères à plusieurs tours. L'enchère à un tour qui a suscité le plus d'intérêt dans la littérature est l'enchère Vickrey-Clarke-Groves (VCG). Comparées aux enchères à un tour, les enchères à plusieurs tours ont l'avantage de révéler graduellement les intentions de chaque participant, permettant ainsi aux enchérisseurs d'ajuster progressivement leurs stratégies en conséquence. Les deux mécanismes d'enchères à plusieurs tours les plus populaires pour la vente du spectre licencié sont l'enchère ascendante simultanée (SAA) et les enchères combinatoires au cadran (CCA). Bien qu'ayant été largement utilisé pour l'allocation de la 4G, le CCA a été délaissé pour l'allocation de la 5G au profit du SAA. Pourquoi ce retour soudain à l'utilisation du SAA ? Malgré certains avantages théoriques du CCA, ses règles sont complexes et opaques. De nos jours, on constate que les régulateurs préfèrent privilégier la simplicité des règles plutôt que certaines garanties théoriques. Par conséquent, il semble que le SAA continuera d'être le mécanisme privilégié pour les enchères de spectre licencié, du moins dans un avenir proche.

## Enchérir dans l'enchère ascendante simultanée (SAA)

Comme décrit dans l'introduction, il est crucial pour les opérateurs mobiles de développer une stratégie d'enchérissement efficace, car l'acquisition des licences a un impact



significatif sur leurs plans d'affaires, et certaines licences représentent des investissements conséquents. L'objectif de ce chapitre est de décrire formellement le problème de l'enchérisseur.

## Mécanisme

Le SAA est un mécanisme où  $m$  objets indivisibles sont vendus simultanément via des enchères anglaises. L'enchère se déroule en plusieurs tours. À chaque tour, les joueurs soumettent leurs offres simultanément. Le joueur ayant soumis la plus grande offre sur un objet le remporte temporairement. Si plusieurs joueurs soumettent la même plus grande offre sur le même objet, le gagnant temporaire est tiré au sort parmi ces joueurs. Le prix courant de l'objet  $j$ , noté  $P_j$ , est ensuite fixé à la plus grande offre soumise pour cet objet. Le prix courant et le gagnant temporaire de chaque objet sont révélés à la fin de chaque tour. L'offre minimale admissible qu'un joueur peut soumettre au prochain tour sur un objet  $j$  est égale à  $P_j + \varepsilon$ , avec  $\varepsilon$  représentant l'incrément des prix. L'enchère se clôture si aucune nouvelle offre n'a été soumise pendant un tour. À la fin de l'enchère, les objets sont vendus à leurs prix courants aux gagnants respectifs.

De plus, une règle d'activité est parfois introduite au sein du SAA pour sanctionner les joueurs qui ne maintiennent pas un certain niveau d'activité tout au long de l'enchère. Cette règle s'appuie sur une quantité appelée éligibilité, qui détermine sur quelles combinaisons d'objets un joueur a le droit de miser. Si un joueur ne respecte pas la règle d'activité, son éligibilité est réduite, l'empêchant ainsi de miser sur certaines combinaisons d'objets pour le reste de l'enchère. Au cours de cette thèse, nous choisissons la règle d'activité suivante fréquemment utilisée dans la littérature au vue de sa simplicité: le nombre d'objets pour lequel un joueur participe (gagne temporairement ou mise) ne peut jamais augmenter.

Dans les trois modèles du SAA considérés dans cette thèse, le prix de chaque objet au début de l'enchère est fixé à 0. Nous contraignons les nouvelles mises sur un objet  $j$  à  $P_j + \varepsilon$ . Cette réduction de l'espace d'action est très courante dans la littérature sur le SAA. De plus, nous faisons l'hypothèse classique qu'un joueur ne misera pas sur un objet qu'il a temporairement remporté.

## Caractéristiques des joueurs

Chaque joueur  $i$  est défini par trois caractéristiques: sa fonction de valorisation  $v_i$ , son budget  $b_i$  et son éligibilité  $e_i$ . Les deux premières caractéristiques sont des données privées et la dernière est publique. Sans perte de généralité,  $v_i$  et  $b_i$  sont choisis indépendamment. Si un joueur  $i$  a temporairement remporté un ensemble d'objets  $Y$ , a une éligibilité de  $e_i$  et que le vecteur des prix courants des objets est  $P$ , alors il ne pourra miser sur un ensemble d'objets  $X$  si et seulement si:

$$\begin{cases} |X| + |Y| \leq e_i \\ \sum_{j \in X} (P_j + \varepsilon) \leq b_i - \sum_{j \in Y} P_j \end{cases}$$

La première contrainte est due à la règle d'activité (contrainte d'éligibilité). La deuxième contrainte vient du fait que l'on suppose qu'un joueur ne misera jamais au dessus de son budget (contrainte budgétaire).

À la fin de l'enchère, l'utilité obtenue par le joueur  $i$  après avoir remporté l'ensemble d'objets  $X$  au vecteur des prix courants  $P$  est égale à:

$$\sigma_i(X, P) = v_i(X) - \sum_{j \in X} P_j$$

Nous supposons que les fonctions de valorisation sont normalisées ( $v_i(\emptyset) = 0$ ), finies et vérifient la condition de libre disposition, c'est-à-dire que pour deux ensembles d'objets  $X$  et  $Y$  si  $X \subset Y$  alors  $v_i(X) \leq v_i(Y)$ .

## Complexité du jeu induit par SAA

Le jeu induit par SAA a les propriétés générales de jeu suivantes: N-joueurs, à somme non constante (la somme des utilités des joueurs différent selon le profil de stratégies), simultané (les joueurs soumettent leurs offres simultanément), non-déterministe (un tirage aléatoire a lieu en cas d'égalité des plus grandes mises sur un objet), à information imparfaite (les joueurs ne savent pas ce que leurs adversaires ont joué pendant le même tour d'enchère) et à information incomplète (les fonctions de valorisations et les budgets des adversaires ne sont pas connus). Cela en fait un jeu déjà très difficile à étudier. De plus, ce jeu a

une très grande complexité. Pour mesurer la complexité d'un jeu, nous utilisons deux métriques dans cette thèse: la *complexité de l'espace des ensembles d'information* et la *complexité de l'arbre de jeu*. Un ensemble d'information est un ensemble d'états du jeu indiscernable par le joueur à jouer, c'est à dire que le joueur n'a pas l'information nécessaire pour discerner les états du jeu différents appartenant à un même ensemble d'information. Sachant qu'une stratégie est une fonction associant à chaque ensemble d'information une action, la *complexité de l'espace des ensembles d'information* correspond à la taille de l'ensemble de départ d'une stratégie. La *complexité de l'arbre de jeu* correspond au nombre de séquences d'actions différentes dans le jeu. Pour donner une idée de l'ordre de grandeur des deux complexités, prenons l'exemple de l'enchère SAA qui a eu lieu en Italie en 2018 où 12 licences de spectre 5G ont été vendues entre 5 entreprises de télécommunications après 171 tours. Dans sa version à information complète considérée au chapitre 5, la complexité de l'espace des ensembles d'informations et la complexité de l'arbre de jeu sont respectivement supérieurs à  $10^{35}$  et  $10^{2470}$ .

En plus des difficultés liées aux propriétés générales du jeu et à la complexité très élevée du SAA, plusieurs problèmes stratégiques s'ajoutent en raison de ses règles spécifiques et de son mécanisme particulier. Les quatre problèmes stratégiques majeures du SAA sont le problème d'exposition, la montée des prix, les contraintes budgétaires et la gestion d'éligibilité. Le problème d'exposition survient lorsqu'un joueur tente d'acquérir des objets complémentaires mais finit par payer trop cher pour ceux qu'il obtient réellement, ce qui le conduit à une utilité négative. Prenons l'exemple de la vente de 2 objets par SAA entre 2 joueurs. Le premier joueur veut uniquement obtenir l'objet 1 et lui donne une valeur de 30, i.e.,  $v_1(\{1\}) = v_1(\{1, 2\}) = 30$  et  $v_1(\{2\}) = 0$ . Le second joueur voit les 2 objets comme des parfaits compléments et leur donne une valeur de 20, i.e.,  $v_2(\{1\}) = v_2(\{2\}) = 0$  et  $v_2(\{1, 2\}) = 20$ . Dans ce cas, si le joueur 2 enchérit sur les 2 objets, il finira exposé car soit il remportera qu'un seul objet soit il devra payer un prix supérieur à 30 pour l'objet 1. Dans les deux cas, il finira avec une utilité négative. La montée des prix vient du fait que chaque mise sur un objet  $j$  augment son prix courant  $P_j$  et donc diminue l'utilité de tous les joueurs voulant l'acquérir. Tous les joueurs ont donc mutuellement intérêt à garder les prix bas. Pour éviter cette augmentation des prix, les joueurs peuvent réduire leurs demandes ou former des coalitions de manière implicite avec les autres joueurs. Les contraintes budgétaires peuvent empêcher les joueurs de miser sur certaines combinaisons d'objets, constituant ainsi une source potentielle d'exposition. Gérer efficacement son éligibilité est un facteur clé pour garantir un résultat favorable. Enchérir sur de nombreux objets permet de maintenir une éligibilité élevée, mais cela induit une augmentation des prix. Cependant, réduire son éligibilité pour former des coalitions peut piéger un joueur dans une position

vulnérable si les autres joueurs ne se comportent pas comme prévu. Il est donc nécessaire de trouver un compromis.

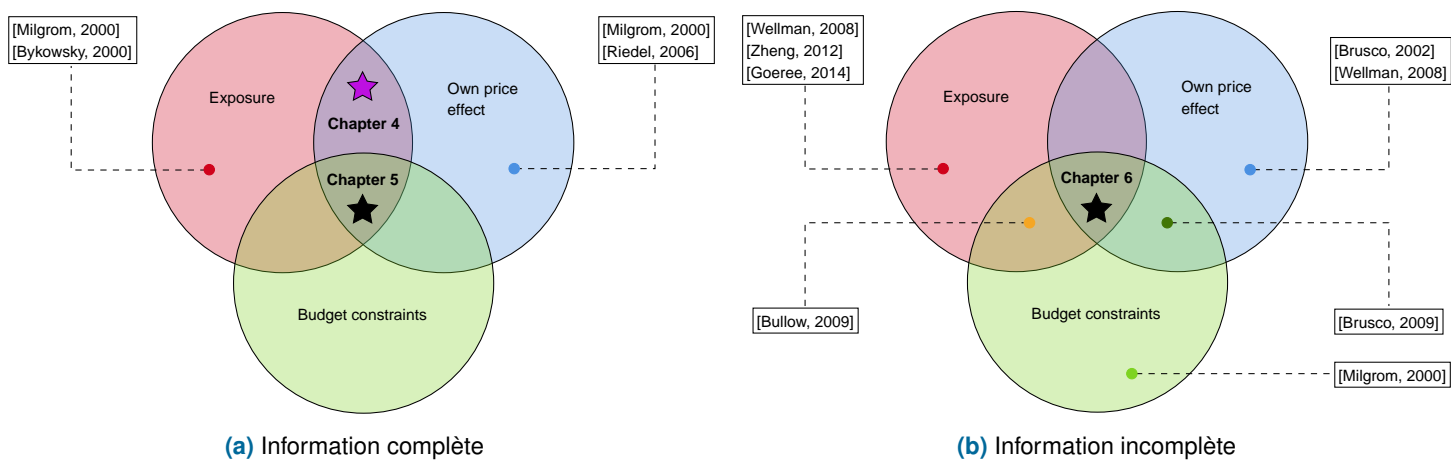
## Indicateurs de performance

La métrique naturelle pour évaluer la performance d'une stratégie est *l'utilité espérée*. Cependant, puisqu'une instance spécifique d'une enchère du spectre licencié (mêmes bandes de fréquences, mêmes opérateurs, etc.) a généralement lieu qu'une seule fois et qu'un opérateur ne participe qu'à un nombre limité d'instances différentes, la comparaison des stratégies uniquement sur la base de leur utilité espérée n'est pas suffisante. En raison des montants d'argent investis, les pertes potentielles dues à l'exposition doivent également être prises en compte. Pour ce faire, nous introduisons deux métriques: *l'exposition espérée*, qui est l'opposé de la somme des pertes d'une stratégie divisée par le nombre de fois qu'elle a été utilisée, et la *fréquence d'exposition*, qui est le nombre de fois qu'une stratégie a été exposée divisé par le nombre de fois qu'elle a été utilisée. Pour augmenter son utilité, un joueur peut soit essayer d'acquérir un ensemble d'objets de plus grande valeur, soit diminuer le prix payé par objet remporté. Afin de mettre en évidence l'augmentation de l'utilité espérée résultant d'une meilleure gestion de la montée des prix, nous examinons le *prix moyen payé par objet remporté*. Pour garantir qu'une stratégie répartit efficacement les objets entre les joueurs et qu'aucun objet ne reste invendu inutilement, nous considérons le *ratio d'objets remportés*.

## Positionnement de la thèse

Dans la Figure 1, nous présentons comment cette thèse s'insère dans la littérature. Nous représentons graphiquement les problèmes stratégiques majeures abordés à la fois par la littérature et par les différents chapitres de cette thèse. La gestion de l'éligibilité n'est pas incluse dans la représentation, car cette question n'a pas été véritablement traitée dans la littérature. Cependant, elle sera abordée dans les chapitres 5 et 6 de cette thèse. Par ailleurs, nous faisons une distinction entre les enchères avec information complète et celles avec information incomplète. Dans le chapitre 4, nous présentons le premier algorithme performant qui traite simultanément le problème d'exposition et la montée des prix dans de petites instances d'une version déterministe de SAA en tour par tour avec information complète. Le chapitre 5 introduit le premier algorithme performant qui aborde simultanément les quatre problématiques stratégiques majeures du SAA dans son format

original à information complète et sur des instances de taille réelle. Enfin, dans le chapitre 6, nous étendons l'algorithme du chapitre précédent au SAA avec information incomplète.



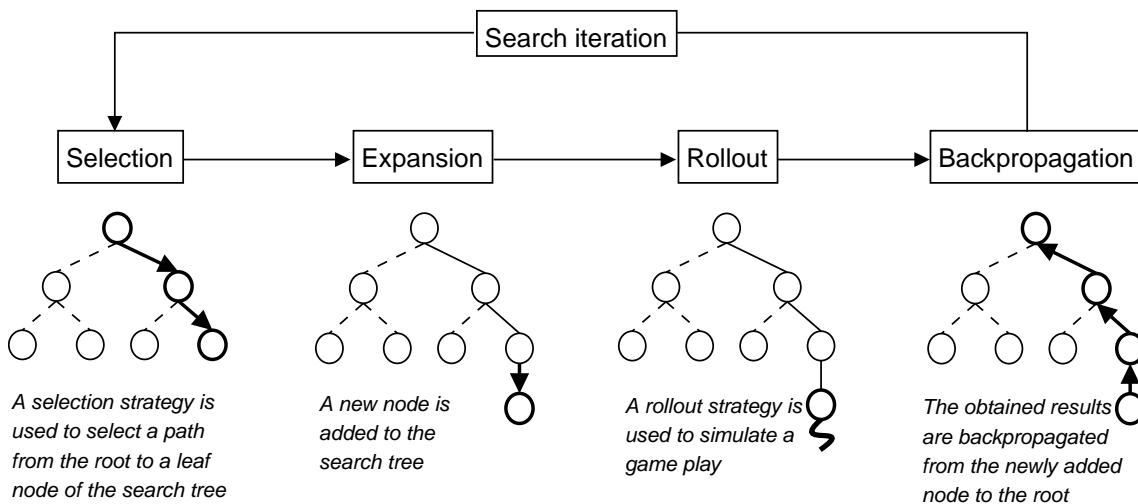
**Fig. 1.:** Les problèmes stratégiques du SAA abordés par la littérature et par nos différents chapitres

## Une revue de la littérature sur les méthodes de recherche dans les jeux avec des adversaires

Dans cette thèse, notre objectif est de concevoir une stratégie performante pour enchérir dans le SAA. Étant donné que le SAA peut être modélisé comme un jeu à plusieurs tours impliquant de nombreux joueurs aux intérêts potentiellement conflictuels, il est essentiel d'intégrer la réaction des autres joueurs dans le développement de notre stratégie. Ainsi, notre attention se porte particulièrement sur les méthodes de recherche dans les jeux avec des adversaires. Dans le chapitre 3, nous exposons diverses méthodes utilisées au fil des années, telles que l'algorithme minimax, l'élagage alpha-bêta, et les bandits. Nous mettons particulièrement l'accent sur le MCTS, qui sera la méthode sur lequel s'appuie principalement tous les algorithmes de cette thèse.

L'arbre de jeu du SAA étant trop grand, il est impossible de l'explorer intégralement. Nous ne pouvons explorer qu'une petite partie appelée arbre de recherche. Le MCTS est un algorithme de parcours en largeur qui construit de manière itérative cet arbre de recherche. Chaque nœud de l'arbre de recherche correspond à un état du jeu et les arêtes aux différentes actions pouvant être effectuées par les joueurs. La figure 2 illustre le processus itératif du MCTS. Chaque itération de recherche est divisée en quatre phases: (1) La phase de sélection consiste à sélectionner un chemin de la racine à un nœud feuille de l'arbre de recherche en utilisant une stratégie de sélection. Un nœud feuille de l'arbre de recherche

est soit un nœud terminal du jeu, soit un nœud dont tous les enfants ne sont pas inclus dans l'arbre de recherche. (2) La phase d'expansion consiste à ajouter un ou plusieurs enfants du nœud feuille sélectionné dans l'arbre de recherche. (3) La phase de simulation simule une partie du jeu jusqu'à sa conclusion à partir du nœud nouvellement ajouté. (4) La phase de rétropropagation consiste à utiliser les résultats de la simulation pour mettre à jour les différentes statistiques stockées dans chaque nœud sélectionné.

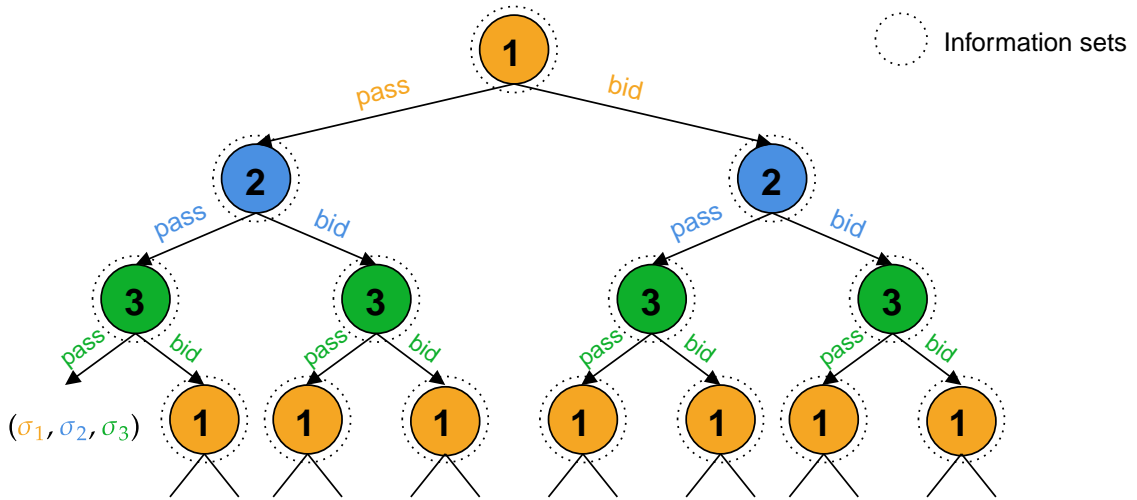


**Fig. 2.:** Représentation graphique du processus itératif de la recherche arborescente de Monte Carlo

## Enchère ascendante simultanée déterministe au tour par tour à information complète

Dans ce chapitre, nous considérons une version simplifiée du SAA, que nous appelons d-SAA. La seule différence par rapport au mécanisme original du SAA réside dans le fait que les joueurs participent de manière séquentielle. À chaque tour, un joueur unique soumet une offre qui est immédiatement révélée aux autres participants. Ce joueur devient temporairement le gagnant des objets pour lesquels son offre est admissible. Ce changement de mécanisme vise à éliminer la stochasticité et la simultanéité de notre problème. De plus, étant donné que les joueurs ont une connaissance parfaite de toutes les actions passées, le jeu est à information parfaite, et donc les ensembles d'information sont des singletons. Les joueurs ne sont pas soumis à des contraintes budgétaires ni aux règles d'activité. L'enchère prend fin lorsqu'aucune offre admissible n'a été soumise pendant  $n$  tours. Le jeu induit par d-SAA est un jeu séquentiel déterministe à  $N$  joueurs

avec une information parfaite et complète. Nous représentons sa forme extensive dans le cadre d'une enchère à trois joueurs et un item dans la figure 3.



**Fig. 3.:** Le jeu induit par d-SAA sous forme extensive avec trois joueurs et un item

## Prédiction des prix finaux

La phase de simulation de notre MCTS est basée sur une stratégie développée par Wellman et al. (2008), que l'on appelle *PP*, et sur une prédiction des prix finaux particulière. La stratégie *PP* consiste à enchérir sur un ensemble d'objets qui maximise l'utilité du joueur selon sa prédiction des prix finaux.

**Definition.** Avec une prédiction initiale des prix finaux  $P_{init}$  et un vecteur des prix courants des objets  $P$ , un enchérisseur qui a temporairement gagné un ensemble d'objets  $Y$  et qui applique la stratégie *PP* a comme prédiction des prix finaux:

$$\forall j \in \{1, \dots, m\}, \rho_j(P^{init}, P, Y) = \begin{cases} \max(P_j^{init}, P_j) & \text{if } j \in Y \\ \max(P_j^{init}, P_j + \varepsilon) & \text{otherwise} \end{cases}$$

Le joueur mise alors sur l'ensemble d'objets  $X^*$ :

$$X^* = \arg \max_{X \subset \{1, \dots, m\} \setminus Y} \sigma(X \cup Y, \rho(P^{init}, P, Y))$$

Dans la littérature, plusieurs méthodes ont été proposées pour calculer une prédiction des prix finaux. Cependant, aucune de ces méthodes ne semble satisfaire ces trois critères : (1) Elle doit être applicable à toutes les instances d'enchères. (2) Elle doit prendre en compte les particularités de chaque enchère. (3) La prédiction finale ne doit pas dépendre d'un profil de stratégies spécifique. Ainsi, nous présentons ci-dessous une nouvelle méthode de prédiction des prix finaux qui répond à ces trois propriétés.

**Definition.** Pour toute instance  $\Gamma$  de d-SAA, on définit  $f_\Gamma$  comme la fonction associant à tout vecteur de prix  $p \in \mathbb{R}_+^m$  les prix finaux  $f_\Gamma(p)$  obtenus dans  $\Gamma$  lorsque tous les enchérisseurs jouent la stratégie PP en utilisant comme prédiction initiale des prix finaux  $p$ .

Théoriquement, une expression analytique de  $f_\Gamma$  peut être calculée pour n'importe quelle instance  $\Gamma$  de d-SAA. Cependant, en pratique, cela n'est possible que pour de petites instances. Ainsi,  $f_\Gamma(p)$  est généralement calculée numériquement en simulant un d-SAA où tous les joueurs jouent selon la stratégie PP avec une prédiction initiale  $p$ .

**Conjecture.** Pour toute instance  $\Gamma$  de d-SAA, la suite  $P_{t+1} = \frac{1}{t+1}f_\Gamma(P_t) + (1 - \frac{1}{t+1})P_t$  avec  $P_0$  le vecteur nul des prix courants converge vers un unique élément  $P^*$ .

Nous observons que la suite  $P_t$  converge pour toutes les instances de d-SAA que nous avons considéré au cours de cette thèse et nous utilisons sa limite  $P^*$  comme prédiction des prix finaux pour notre MCTS.

## Stratégie d'enchérissement $MS^\lambda$

Nous implémentons une variante de MCTS-max<sup>n</sup>, que l'on appelle  $MS^\lambda$ . Chaque nœud  $x$  stocke les caractéristiques suivantes: la somme des récompenses  $r_x$  obtenus dans le sous-arbre correspondant, le nombre de visites  $n_x$ , la borne inférieure estimée  $a_x$  et la borne supérieure estimée  $c_x$  du support des récompenses. Les quatre phases du MCTS sont décrites ci-dessous.



- **Sélection:** Nous utilisons comme stratégie de sélection un UCT (*Upper Confidence Bounds applied to Trees*) avec des pénalités. À partir d'un nœud sélectionné  $y$ , la stratégie de sélection choisit l'enfant  $x$  ayant le score  $q_x$  le plus élevé:

$$q_x = \frac{r_x}{n_x} + \max(c_x - a_x, \varepsilon) \sqrt{\frac{2 \log(n_y)}{n_x}} - no\_object(x) - risky\_move(x)$$

Les deux pénalités de sélection  $no\_object(x)$  et  $risky\_move(x)$  sont introduites pour réduire le problème d'exposition.

- **Expansion:** Un enfant du nœud feuille est ajouté aléatoirement à l'arbre. Ses statistiques sont initialisées de la manière suivante:  $r_x = 0$ ,  $n_x = 0$ ,  $a_x = +\infty$  et  $c_x = -\infty$ .
- **Simulation:** À partir du nœud nouvellement ajouté, une partie de d-SAA est simulée jusqu'à sa clôture, où tous les joueurs jouent selon la stratégie PP avec une prédiction initiale  $P^{init} = P^* + \eta$ , où  $\eta$  est une variable aléatoire suivant une loi uniforme  $U([- \varepsilon, \varepsilon])$ .
- **Rétropropagation:** Soit  $V$  le vecteur d'utilité obtenu lors de la phase de simulation. Soit  $x$  un nœud sélectionné,  $y$  son nœud parent et  $i$  le joueur jouant au nœud  $y$ . Les statistiques stockées au nœud  $x$  sont mises à jour de la façon suivante:  $r_x \leftarrow r_x + V_i$ ,  $n_x \leftarrow n_x + 1$ ,  $a_x \leftarrow \min(a_x, V_i)$  et  $c_x \leftarrow \max(c_x, V_i)$ .

Lorsque toutes les itérations de recherche ont été réalisées,  $MS^\lambda$  renvoie l'action conduisant à l'enfant  $x$  du nœud racine avec le score  $q_x$  le plus élevé:

$$q_x = \frac{r_x}{n_x} - no\_object(x) - risky\_move(x)$$

## Expériences numériques

Nous évaluons les performances de la stratégie  $MS^\lambda$  sur des petites instances ( $n = 2$  et  $m = 7$ ) en la comparant à cinq autres stratégies:  $MS^{np}$ , similaire à  $MS^\lambda$  mais sans les pénalités de sélection; un algorithme UCB; la stratégie SB, équivalente à une stratégie PP avec une prédiction initiale nulle; la stratégie EPE, une stratégie PP avec une prédiction

initiale particulière; et la stratégie SCPD, très similaire à une stratégie PP mais utilisant une distribution des prix finaux. Chaque indicateur de performance est calculé pour chaque profil de stratégies à partir de 1000 instances de d-SAA où les fonctions de valorisation ont été générées aléatoirement. Pour faciliter l'analyse, nous étudions le jeu sous forme normale en utilité espérée où un joueur doit choisir entre jouer la stratégie  $MS^\lambda$  ou une autre stratégie  $A$ . Les 5 jeux empiriques possibles sont illustrés dans la figure 4. Dans cette figure, nous voyons que la déviation d'une stratégie  $A \in \{UCB, SB, EPE, SCPD, MS^{np}\}$  vers  $MS^\lambda$  est toujours profitable. De ce fait, le profil de stratégies  $(MS^\lambda, MS^\lambda)$  est un équilibre de Nash du jeu d-SAA avec ces six stratégies. Par ailleurs, nous montrons à travers les autres indicateurs de performance que la montée des prix et le problème d'exposition sont significativement mieux traités avec  $MS^\lambda$  qu'avec une des cinq autres stratégies.

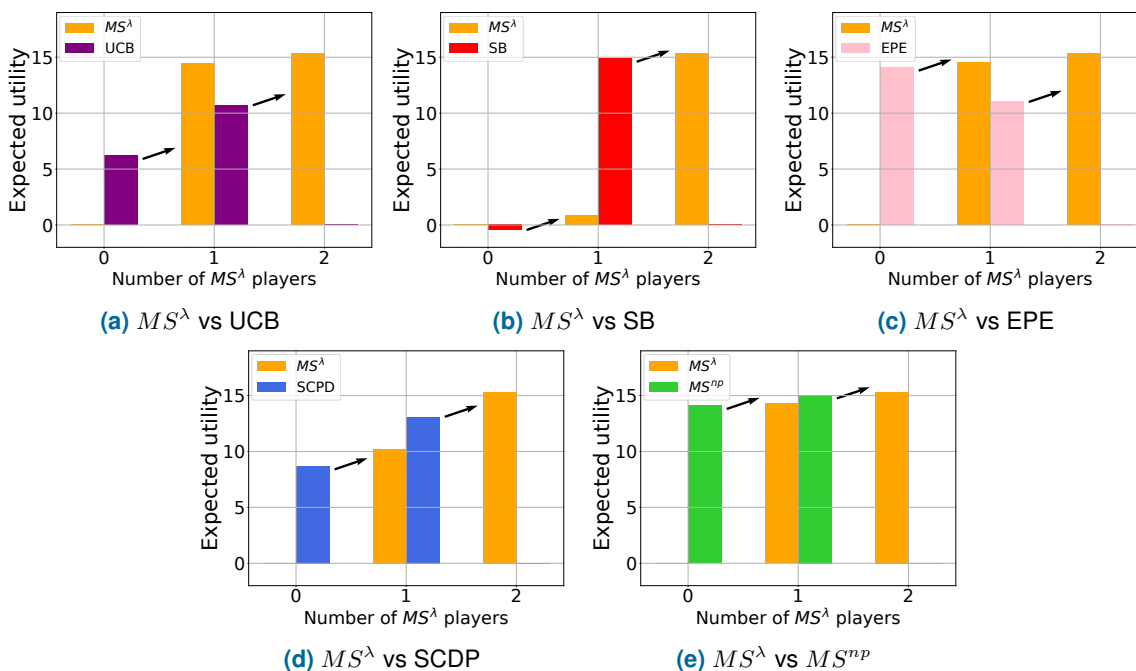
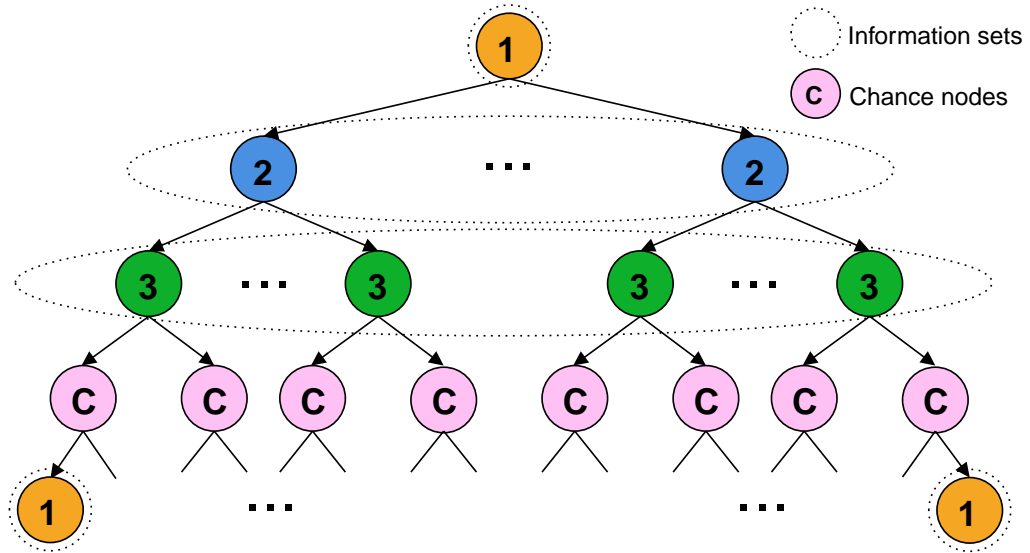


Fig. 4.: Le jeu d-SAA sous forme normale avec six stratégies

## Enchère ascendante simultanée à information complète

Dans ce chapitre, nous considérons le format original du SAA à information complète, que nous appelons SAA-c. À la différence du chapitre précédent, les joueurs sont désormais soumis à des règles d'activité et des contraintes budgétaires. Étant donné que les offres

sont soumises simultanément, chaque joueur ne peut connaître les mises des autres joueurs pendant le tour en cours. Par conséquent, les états du jeu qui diffèrent uniquement par les mises effectuées au cours d'un même tour appartiennent au même ensemble d'information. Le jeu induit par SAA-c est un jeu simultané non-déterministe à N joueurs à information imparfaite et complète. Sa forme extensive est illustrée dans la figure 5, avec ses ensembles d'information et ses nœuds chance représentant le tirage aléatoire du gagnant temporaire en cas d'égalité de la plus grande mise soumise sur un objet.



**Fig. 5.:** Le jeu induit par SAA-c avec trois joueurs sous forme extensive avec ses ensembles d'information et ses nœuds chances

## Prédiction des prix finaux

Nous étendons la stratégie PP, définie au chapitre précédent, aux contraintes budgétaires et d'éligibilité, c'est-à-dire qu'un joueur appliquant la stratégie PP mise sur l'ensemble d'objets  $X^*$ :

$$X^* = \underset{\substack{X \subset \{1, \dots, m\} \setminus Y \\ \sum_{j \in X \cup Y} \rho_j(P^{init}, P, Y) \leq b \\ |X| + |Y| \leq e}}{\arg \max} \sigma(X \cup Y, \rho(P^{init}, P, Y))$$

avec  $P^{init}$  une prédiction initiale des prix finaux,  $P$  le vecteur des prix courants,  $Y$  l'ensemble d'objets temporairement gagné par le joueur,  $e$  son éligibilité et  $b$  son budget.

Avec cette extension de la stratégie PP, nous adaptons notre méthode de prédiction des prix finaux du chapitre précédent aux contraintes budgétaires, aux contraintes d'éligibilité et à la stochasticité du SAA-c. Cette méthode est aussi fondée sur la convergence d'une suite spécifique qui vise à satisfaire les trois mêmes critères.

**Conjecture.** Soit  $\Gamma$  une instance du jeu induit par SAA-c. Soit  $f_\Gamma(P)$  la variable aléatoire renvoyant les prix finaux de  $\Gamma$  lorsque tous les joueurs jouent selon la stratégie PP avec une prédiction initiale  $P$ . La suite  $p_{t+1} = \frac{1}{t+1} \mathbb{E}[f_\Gamma(p_t)] + (1 - \frac{1}{t+1})p_t$  avec  $p_0$  le vecteur nul des prix converge vers un unique point  $p^*$ .

Le fait que  $f_\Gamma$  soit une variable aléatoire provient du tirage aléatoire en cas d'égalité des plus grandes mises sur un objet. Dans la pratique, nous procédons à une estimation Monte-Carlo de  $\mathbb{E}[f_\Gamma(p_t)]$  en simulant l'enchère de nombreuses fois. Cependant, il est également possible d'obtenir une expression analytique de  $\mathbb{E}[f_\Gamma(p_t)]$  et, à partir de là, de démontrer la convergence de la suite  $p_t$ . La limite conjecturée  $p^*$  de cette suite est utilisée par l'algorithme MCTS décrit dans la section suivante.

## Stratégie d'enchérissement $SMS^\alpha$

Dans cette section, nous présentons un nouvel algorithme MCTS, que nous appelons  $SMS^\alpha$ , qui s'appuie sur une fonction d'utilité particulière. Nous l'appelons "utilité scalarisée" et elle a pour objectif de maximiser l'utilité d'un joueur tout en minimisant les risques d'exposition.

$$\sigma^\alpha(X, P) = (1 + \alpha \mathbb{1}_{\sigma(X, P) < 0}) \sigma(X, P)$$

Comparées à l'utilité précédemment définie, les pertes liées au problème d'exposition sont plus sévèrement pénalisées. L'hyperparamètre  $\alpha$  permet d'arbitrer entre utilité et aversion au risque.

Nous présentons ci-dessous les différentes phases de notre algorithme  $SMS^\alpha$ :

- **Sélection:** À chaque étape de sélection, nous choisissons l'action qui maximise l'indice de sélection pour chaque joueur  $i$ . Les actions des joueurs sont ainsi sélectionnées simultanément et indépendamment les unes des autres. Chaque étape de

sélection correspond donc à un tour entier de SAA-c. Notre indice de sélection est une application directe de UCT aux utilités scalarisées. Le joueur  $i$  choisit de miser sur l'ensemble d'objets  $x_i$  avec le score  $q_{x_i}$  le plus élevé à l'ensemble d'information  $I_i$ :

$$q_{x_i} = \frac{r_{x_i}^\alpha}{n_{x_i}} + \max(c_{x_i}^\alpha - a_{x_i}^\alpha, \varepsilon) \sqrt{\frac{2 \log(\sum_{x'_i} n_{x'_i})}{n_{x_i}}} \quad (0.1)$$

avec  $r_{x_i}^\alpha$  la somme des utilités scalarisées obtenues après avoir enchéri sur  $x_i$  à  $I_i$ ,  $n_{x_i}$  le nombre de fois que le joueur  $i$  a misé sur  $x_i$  à  $I_i$ , la borne inférieure estimée  $a_{x_i}^\alpha$  et la borne supérieure estimée  $c_{x_i}^\alpha$  du support des utilités scalarisées.

- **Expansion:** Le facteur de branchement élevé du jeu induit par SAA-c empêche d'inspecter de manière approfondie les branches prometteuses de l'arbre de recherche et donc de planifier une stratégie sur de nombreux tours. Par conséquent, il est nécessaire de réduire l'espace d'actions à chaque ensemble d'information de l'arbre de recherche. Nous limitons le nombre maximum d'actions pour chaque ensemble d'information inclus dans l'arbre de recherche à  $N_{act}$ . Miser sur aucun objet est toujours compris dans les  $N_{act}$  actions. Les  $N_{act} - 1$  actions restantes correspondent à celles avec la plus grande utilité selon la stratégie PP avec une prédiction initiale de  $p^*$ .
- **Simulation:** Au début de chaque phase de simulation, nous fixons, pour chaque joueur  $i$ ,  $p_i^* = p^* + \eta_i$  avec  $\eta_i \sim U([- \varepsilon, \varepsilon])$ . Chaque joueur  $i$  joue alors selon la stratégie PP avec sa propre prédiction initiale  $p_i^*$ . Nous ajoutons du bruit à notre prédiction initiale  $p^*$  afin de diversifier les stratégies des joueurs dans la phase de simulation et améliorer la qualité de l'échantillonnage.
- **Rétropropagation:** Les différentes statistiques sont mises à jour en utilisant l'utilité scalarisée de chaque joueur obtenue lors de la phase de simulation.

Lorsque toutes les itérations de recherche ont été réalisées,  $SMS^\alpha$  renvoie l'action qui maximise notre estimation de l'utilité scalarisée espérée du joueur racine.

## Expériences numériques

Nous évaluons la performance de  $SMS^\alpha$  sur des instances de taille réelle ( $n = 4$  et  $m = 11$ ) en la comparant à quatre stratégies de l'état de l'art:  $MS^\lambda$ , EPE, SCPD et SB. Les

fonctions de valorisation et budgets des joueurs sont générés aléatoirement. Pour faciliter notre analyse, nous étudions le jeu sous forme normale en utilité espérée, où chaque joueur doit choisir entre jouer  $SMS^\alpha$  ou une autre stratégie  $A$ . L'utilité espérée de chaque profil de stratégies est calculée à partir de 1000 instances de SAA-c. Nous représentons dans la figure 6 les quatre jeux empiriques possibles. Dans chaque jeu empirique, un joueur a toujours intérêt à jouer  $SMS^\alpha$  pour augmenter son utilité. Par conséquent, le profil de stratégies  $(SMS^\alpha, SMS^\alpha, SMS^\alpha, SMS^\alpha)$  est un équilibre de Nash du jeu SAA-c avec ces cinq stratégies. De plus, nos autres indicateurs de performance révèlent que le problème d'exposition et la montée des prix est significativement mieux traité avec  $SMS^\alpha$  en présence de contraintes budgétaires et d'éligibilité.

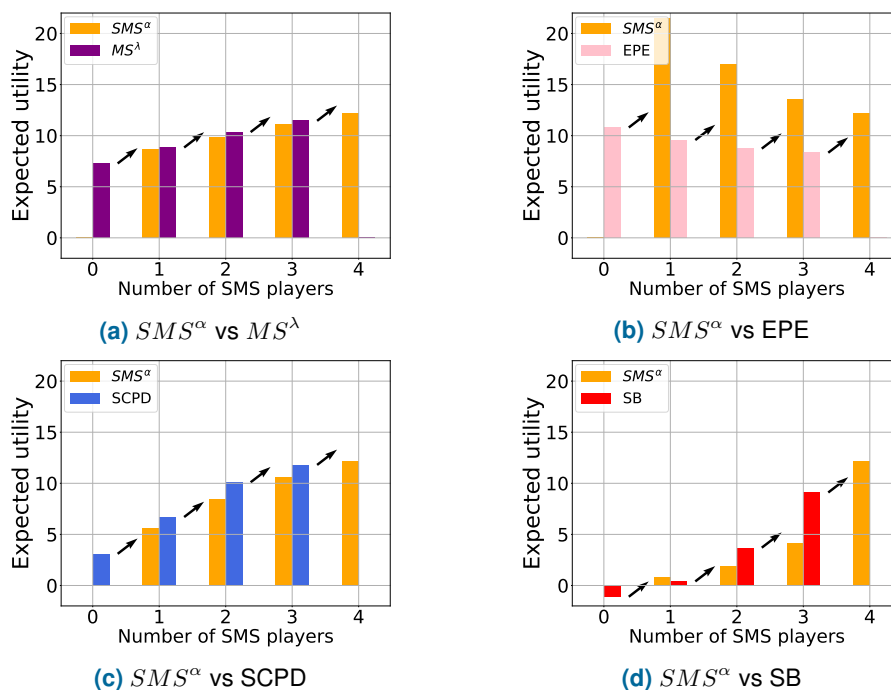
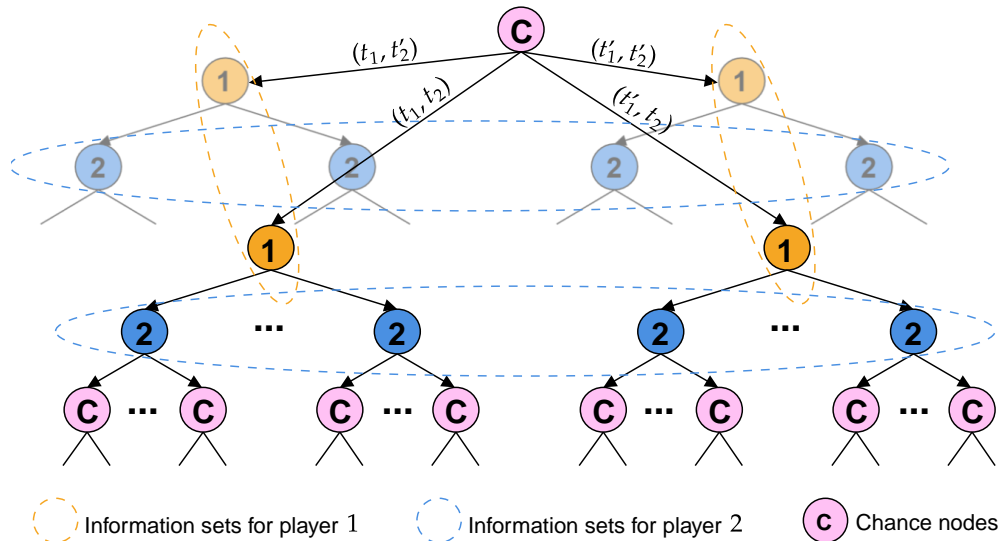


Fig. 6.: Le jeu SAA-c sous forme normale en utilité espérée avec cinq stratégies

## Enchère ascendante simultanée à information incomplète

Le chapitre précédent présentait une approche pour enchérir dans un jeu d'enchères à information complète. Dans ce chapitre, nous relâchons cette hypothèse et considérons le même jeu d'enchères, mais avec une information incomplète. Les joueurs n'ont donc plus une estimation parfaite des fonctions de valorisation et des budgets de leurs adversaires. Nous appelons ce modèle SAA-inc.

Afin de modéliser l'incertitude des joueurs sur l'information privée de leurs adversaires, nous nous repons sur la notion de *type*. Un type est une combinaison possible de l'information privée d'un joueur, c'est à dire une fonction de valorisation et un budget. Par exemple, si le joueur 1 a deux fonctions de valorisation possibles,  $v_1$  et  $v_2$ , ainsi que deux budgets possibles,  $b_1$  et  $b_2$ , du point de vue du joueur 2, alors le joueur 1 présente 4 types possibles:  $(v_1^1, b_1^1)$ ,  $(v_1^1, b_2^1)$ ,  $(v_2^1, b_1^1)$  et  $(v_2^1, b_2^1)$ . Chaque joueur a une estimation probabiliste des types de ses adversaires. Nous supposons que tous les adversaires du joueur  $i$  partagent la même estimation probabiliste de son type. Notons alors  $\mathcal{T}_i$  la distribution du type du joueur  $i$ . Cette hypothèse est fréquemment utilisée dans la littérature. Chaque joueur  $i$  connaît son type mais pas celui de ses adversaires, c'est à dire qu'il connaît  $(v_i, b_i)$  et possède une distribution  $\mathcal{T}_{i'}$  du type de chaque adversaire  $i'$ . Nous représentons une version équivalente, appelée "équivalent de Bayes", du jeu induit par SAA-inc sous forme extensive dans la figure 7. Le nœud chance en haut de l'arbre représente le tirage des types des joueurs selon leurs distributions respectives. Comparé à SAA-c, la taille des ensembles d'information augmente car un joueur est incapable de différencier les états du jeu qui ne diffèrent que par le type de leurs adversaires. Par exemple, dans la figure 7, les deux nœuds correspondant au joueur 1 avec les types  $(t_1, t_2)$  et  $(t_1, t'_2)$  appartiennent au même ensemble d'information.



**Fig. 7.:** Représentation sous forme extensive du jeu SAA-inc avec deux joueurs. Le premier nœud chance correspond au tirage du type de chaque joueur. Par exemple, le couple  $(t_1, t'_2)$  signifie que le premier joueur est de type  $t_1$  et que le deuxième joueur est de type  $t'_2$ .

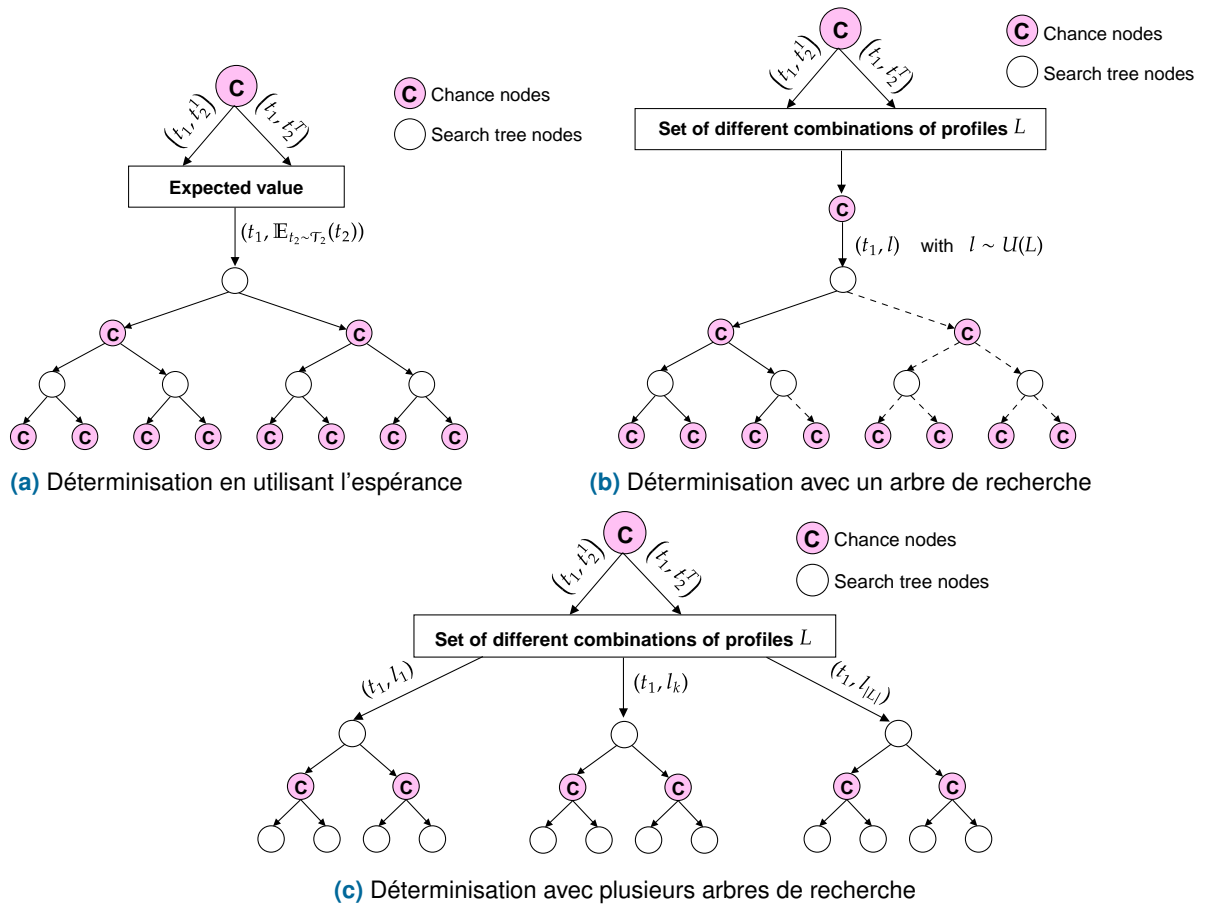
## Trois approches de détermination

La détermination est une méthode permettant de prendre des décisions dans les jeux impliquant de l'information cachée en échantillonnant des instances du jeu équivalent où cette information serait révélée. Dans cette section, nous présentons trois approches de détermination adaptant  $SM S_{EXP3}^\alpha$  (un algorithme similaire à  $SM S^\alpha$  mais où l'indice de sélection UCT est remplacée par EXP3) au cadre d'information incomplète du jeu SAA-inc. La première approche, nommée  $\overline{SM S_{EXP3}^\alpha}$  et représentée dans la figure 8(a), consiste à appliquer  $SM S_{EXP3}^\alpha$  au jeu déterminisé où tous les adversaires jouent selon l'espérance de leur distribution des types. Pour les deux approches suivantes, nous sélectionnons préalablement des combinaisons de fonctions de valorisation et de budgets que nous jugeons pertinentes par rapport à la distribution des types des adversaires. La seconde approche, nommée  $DSM S_{EXP3}^\alpha$  et représentée dans la figure 8(c), consiste à générer un jeu déterminisé pour chaque combinaison et à appliquer  $SM S_{EXP3}^\alpha$  à chaque instance. Parmi toutes les instances, l'action ayant été la plus souvent retournée par  $SM S_{EXP3}^\alpha$  est celle sélectionnée par  $DSM S_{EXP3}^\alpha$ . La troisième approche, nommée  $SDSM S_{EXP3}^\alpha$  et représentée dans la figure 8(b), consiste à générer un unique arbre de recherche et à tirer une détermination différente à chaque itération de recherche. Seules les actions des adversaires cohérentes avec la détermination tirée sont considérées, c'est-à-dire que les actions d'un adversaire ne respectant pas son budget tiré ne sont pas prises en compte.

## Expériences numériques

Tout d'abord, nous comparons la performance de nos trois approches de détermination à un ensemble, appelé PPB, de quatre stratégies de l'état de l'art: EPE, EDPE (stratégie PP avec une prédiction initiale particulière), SCPD et SB. Puis, nous les comparons entre elles. Les expériences sont réalisées pour plusieurs niveaux d'incertitudes sur l'estimation des types des adversaires. Plus le niveau d'incertitude est grand, plus la distribution des types des adversaires sera large. Chaque indicateur de performance est calculé à partir de 1000 instances de SAA-inc. Comme dans le chapitre précédent, nous étudions le jeu sous forme normale en utilité espérée où chaque joueur a le choix entre deux stratégies. Dans chaque jeu empirique où un joueur peut choisir entre une de nos approches de détermination et l'une des quatre stratégies de PPB, nous constatons qu'un joueur a toujours intérêt à jouer l'une de nos approches, quel que soit le niveau d'incertitude. De plus, à travers les autres indicateurs de performance, nous montrons que le problème d'exposition et la montée des prix sont largement mieux traités par nos approches en présence de contraintes budgétaires et d'éligibilité. Les différences entre nos trois approches de détermination





**Fig. 8.:** Représentation des trois approches de détermination appliquées à un jeu SAA-inc avec deux joueurs du point de vue du joueur 1.

sont moins marquées qu'avec les stratégies de PPB. Néanmoins, nous observons que  $SDSM S_{EXP3}^\alpha$  surpasse légèrement  $DSM S_{EXP3}^\alpha$  et  $\overline{SM S_{EXP3}^\alpha}$  en termes de maximisation de l'utilité espérée et de traitement du problème d'exposition. Cependant, si tous les joueurs jouent la même stratégie, jouer  $DSM S_{EXP3}^\alpha$  est plus profitable que jouer  $SDSM S_{EXP3}^\alpha$ .

Malgré ces résultats très prometteurs, nous notons que l'incertitude a un impact significatif sur nos trois approches de détermination, augmentant le risque d'exposition et la montée des prix. Cela se traduit par une diminution de leur utilité espérée respective. De plus, lorsque tous les joueurs adoptent la même stratégie, l'incertitude affecte également la coordination, surtout lorsque la distribution des types est très large. Pour atténuer ce problème, une piste intéressante serait d'utiliser des méthodes d'inférence afin d'affiner la distribution des types des adversaires au fil des tours de SAA-inc.

# Contents

<b>List of Figures</b>	<b>3</b>
<b>List of Tables</b>	<b>5</b>
<b>List of Algorithms</b>	<b>7</b>
<b>Introduction</b>	<b>9</b>
<b>1 An introduction to spectrum auctions</b>	<b>11</b>
1.1 Assigning spectrum licences . . . . .	11
1.2 Spectrum auctions: objectives, properties, issues . . . . .	12
1.3 Main families of spectrum auctions . . . . .	16
<b>2 Bidding in Simultaneous Ascending Auctions</b>	<b>31</b>
2.1 Bidder's characteristics . . . . .	32
2.2 Complexities induced by the bidding game in SAA . . . . .	36
2.3 Performance indicators . . . . .	44
2.4 Benchmarking: perceived-price bidding strategies . . . . .	46
2.5 Other related works . . . . .	55
2.6 Positioning of the thesis . . . . .	60
<b>3 A survey on adversarial search</b>	<b>63</b>
3.1 Extensive form . . . . .	64
3.2 Game tree and search tree . . . . .	67
3.3 Adversarial search in two player zero-sum games . . . . .	68
3.4 Adversarial search in n-player games . . . . .	72
3.5 Adversarial search in games with large state space . . . . .	74
3.6 Bandits . . . . .	76
3.7 Monte Carlo Tree Search . . . . .	79
3.8 Related works on MCTS . . . . .	83
3.9 Transposition table . . . . .	88

3.10 Conclusion . . . . .	90
<b>4 Turn-based deterministic Simultaneous Ascending Auction with complete information</b>	<b>93</b>
4.1 Deterministic model with complete and perfect information of SAA (d-SAA)	94
4.2 Complexities induced by the d-SAA game . . . . .	96
4.3 A new prediction method: frontier prediction of final prices . . . . .	100
4.4 MCTS bidding strategy . . . . .	107
4.5 Experiments . . . . .	113
4.6 Conclusion . . . . .	120
<b>5 Simultaneous Ascending Auction with complete information</b>	<b>121</b>
5.1 Simultaneous Ascending Auction model with constraints and complete information (SAA-c) . . . . .	123
5.2 Complexities induced by the SAA-c game . . . . .	126
5.3 Predicting closing prices . . . . .	128
5.4 SM-MCTS bidding strategy . . . . .	132
5.5 Experiments . . . . .	138
5.6 Conclusion . . . . .	149
<b>6 Simultaneous Ascending Auction with incomplete information</b>	<b>151</b>
6.1 Simultaneous Ascending Auction with constraints and incomplete information (SAA-inc) . . . . .	152
6.2 Complexities induced by the SAA-inc game . . . . .	156
6.3 Integrating EXP3 in $SMS^\alpha$ . . . . .	158
6.4 Three determinization approaches . . . . .	161
6.5 Generating combinations of profiles . . . . .	172
6.6 Tracking bid exposure . . . . .	173
6.7 Type generation . . . . .	174
6.8 Extensive experiments . . . . .	179
6.9 Conclusion . . . . .	197
<b>7 Conclusion</b>	<b>199</b>
7.1 Summary of achievements . . . . .	199
7.2 Short-term prospects . . . . .	202
7.3 Long-term prospects . . . . .	203
<b>Bibliography</b>	<b>205</b>
<b>A Appendix: Example 1 in d-SAA game</b>	<b>221</b>

<b>B Appendix: Proof of properties 4.2 and 4.3</b>	<b>229</b>
<b>C Appendix: Example 1 in SAA-c game</b>	<b>295</b>



# List of Figures

1	Les problèmes stratégiques du SAA abordés par la littérature et par nos différents chapitres . . . . .	xii
2	Représentation graphique du processus itératif de la recherche arborescente de Monte Carlo . . . . .	xiii
3	Le jeu induit par d-SAA sous forme extensive avec trois joueurs et un item . .	xiv
4	Le jeu d-SAA sous forme normale avec six stratégies . . . . .	xvii
5	Le jeu induit par SAA-c avec trois joueurs sous forme extensive avec ses ensembles d'information et ses nœuds chances . . . . .	xviii
6	Le jeu SAA-c sous forme normale en utilité espérée avec cinq stratégies . . .	xxi
7	Représentation sous forme extensive du jeu SAA-inc . . . . .	xxii
8	Représentation des trois approches de déterminisation appliquées à un jeu SAA-inc avec deux joueurs du point de vue du joueur 1. . . . .	xxiv
2.1	Strategical issues addressed by related works and by the different chapters of our thesis in SAA. . . . .	61
3.1	An example of a tree . . . . .	65
3.2	An example of a game in extensive form . . . . .	66
3.3	An example of a minimax search in Tic-Tac-Toe . . . . .	69
3.4	An example of a $\alpha\beta$ pruning . . . . .	71
3.5	An example of $\max^n$ search . . . . .	73
3.6	An example of paranoid search . . . . .	74
3.7	Monte Carlo Tree Search scheme . . . . .	80
4.1	A d-SAA game in extensive form with three players and one item . . . . .	96
4.2	A fictitious example of a cycle in the spatial representation of $f_T$ . . . . .	107
4.3	Evolution of player 1's utility $\sigma_1$ depending on strategy versus player 2's valuation $\ell$ in Test experiment [Brusco, 2002] ( $h = 10, \varepsilon = 0.1$ ) given that player 2 plays optimally . . . . .	115
4.4	Normal-form payoffs for a d-SAA game with six strategies . . . . .	117
4.5	Own price effect analysis of a d-SAA game with six strategies . . . . .	118

4.6	Exposure analysis of a d-SAA game with six strategies . . . . .	119
5.1	Extensive form of a three player SAA-c game with information sets and chance nodes . . . . .	125
5.2	SM-MCTS tree structure with explicit chance nodes for a SAA-c game with 3 players . . . . .	133
5.3	Convergence of sequence $p_t^1$ and $p_t^2$ with their respective upper bound $g(t)$ and lower bound $h(t)$ in the SAA-c game of <i>Example 1</i> . . . . .	139
5.4	Convergence of sequence $p_t^j$ for every item $j$ in an SAA-c game generated from our extensive experiments ( $n = 4, m = 11$ and $\varepsilon = 1$ ) . . . . .	140
5.5	Evolution of player 1's expected utility $\mathbb{E}(\sigma_1)$ depending on strategy versus player 2's budget $b_2$ given that player 2 plays optimally ( $\varepsilon = 0.1$ ). . . . .	141
5.6	Evolution of player 2's expected utility $\mathbb{E}(\sigma_2)$ depending on strategy versus player 1's budget $b_1$ given that player 1 plays optimally in <i>Example 1</i> ( $\varepsilon = 1$ ). . . . .	142
5.7	Normal-form expected utility for a SAA-c game with five strategies . . . . .	145
5.8	Own price effect analysis for a SAA-c game with five strategies . . . . .	146
5.9	Exposure analysis for a SAA-c game with five strategies . . . . .	147
5.10	Impact of $\alpha$ on $SMS^\alpha$ . . . . .	148
6.1	Extensive form of a two player SAA-inc game with information sets and chance nodes . . . . .	156
6.2	Representation of the three different determinization approaches applied to a SAA-inc game with two players from the point of view of player 1. . . . .	162
6.3	Empirical value distributions for obtaining sets $\{1, 2\}$ , $\{1, 2, 3, 4\}$ and $\{1, \dots, m\}$ for level of certainty $\eta_v \in \{0, 0.5, 0.8\}$ by sampling $10^4$ value functions $v_i$ from the value distribution $\mathcal{F}_i$ of player $i$ . The complementarity distribution $\mathcal{G}_i^{\eta_v}$ and value distribution $\mathcal{F}_i$ are built with $V = 5$ and $m = 9$ . . . . .	177
6.4	Evolution of the standard deviation of the empirical value distributions for obtaining sets $\{1, 2\}$ , $\{1, 2, 3, 4\}$ and $\{1, \dots, m\}$ depending on the level of certainty $\eta_v$ . We sample $10^4$ value functions $v_i$ from the value distribution $\mathcal{F}_i$ of player $i$ . The complementarity distribution $\mathcal{G}_i^{\eta_v}$ and value distribution $\mathcal{F}_i$ are built with $V = 5$ and $m = 9$ . . . . .	178
6.5	Comparing approaches based on $SMS_{EXP3}^\alpha$ with PPB approaches through normal-form SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) in expected utility with a level of certainty of $(\eta_v, \eta_b) = (0.5, 0.5)$ . . . . .	184
6.6	Comparing our three determinization approaches through normal-form SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) in expected utility with a level of certainty of $(\eta_v, \eta_b) = (0.5, 0.5)$ . . . . .	186

6.7	Comparing our three determinization approaches to $CSMS_{EXP3}^\alpha$ through normal-form SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) in expected utility with a level of certainty of $(\eta_v, \eta_b) = (0.5, 0.5)$ . . . . .	187
6.8	Comparing the decrease in expected utility (%) obtained when all players play $CSMS_{EXP3}^\alpha$ and when all players play the same strategy $A$ with $A \in \{SDSMS_{EXP3}^\alpha, DSM S_{EXP3}^\alpha, \overline{SMS}_{EXP3}^\alpha\}$ for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with different levels of certainty ( $\eta_v = \eta_b$ ). . . . .	188
6.9	Comparing own price effect of $\overline{SMS}_{EXP3}^\alpha$ with PPB approaches for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with a level of certainty $(\eta_v, \eta_b) = (0.5, 0.5)$ through two performance indicators: the average price paid per item won and the ratio of items won . . . . .	190
6.10	Comparing own price effect of approaches based on $SMS_{EXP3}^\alpha$ for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with a level of certainty $(\eta_v, \eta_b) = (0.5, 0.5)$ through two performance indicators: the average price paid per item won and the ratio of items won . . . . .	191
6.11	Comparing the rise in average price paid per item won (%) when all players play $CSMS_{EXP3}^\alpha$ and when all players play the same strategy $A$ with $A \in \{SDSMS_{EXP3}^\alpha, DSM S_{EXP3}^\alpha, \overline{SMS}_{EXP3}^\alpha\}$ for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with different levels of certainty ( $\eta_v = \eta_b$ ). . . . .	192
6.12	Comparing the exposure suffered by $DSMS_{EXP3}^\alpha$ with the exposure suffered by PPB approaches for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with a level of certainty $(\eta_v, \eta_b) = (0.5, 0.5)$ through two performance indicators: the expected exposure and the exposure frequency . . . . .	194
6.13	Comparing the expected exposure and exposure frequency of our approaches based on $SMS_{EXP3}^\alpha$ for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with a level of certainty $(\eta_v, \eta_b) = (0.5, 0.5)$ . . . . .	195
6.14	Exposure frequency (%) of our three determinization approaches when all players play the same strategy for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with different levels of certainty ( $\eta_v = \eta_b$ ). . . . .	196
A.1	Spatial representation of $f_\Gamma$ and sequence dynamic in the d-SAA game of <i>Example 1</i> . . . . .	222
A.2	Convergence rate of $P_t^1$ and $P_t^2$ in the d-SAA game of <i>Example 1</i> . . . . .	228
C.1	Spatial representation of $f_\Gamma$ in the SAA-c game of <i>Example 1</i> . . . . .	296
C.2	Convergence rate of $p_t^1$ and $p_t^2$ in the SAA-c game of <i>Example 1</i> . . . . .	300





# List of Tables

1.1	The GSM spectrum auction in Germany in 1999. Licences 1 to 9 correspond to bandwidth $2 \times 1$ and licence 10 to $2 \times 1.4$ MHz. Winners are either Mannesmann (M) or T-Mobile (T). . . . .	24
2.1	Example of exposure ( $\varepsilon = 1$ ), also referred to as <i>Example 1</i> . . . . .	41
2.2	Example of own price effect ( $\varepsilon = 1$ ) . . . . .	42
2.3	Example of exposure due to budget constraints ( $\varepsilon = 1, b_1 = 10, b_2 = 10$ ). . . . .	43
3.1	Summary of all MCTS algorithms introduced in this thesis . . . . .	91
5.1	Evolution of the expected utility, the average price per item won and the ratio of allocated items when all bidders play $SM S^\alpha$ and $\alpha$ increases . . . . .	149
6.1	Summary of the features of the SAA-inc instances employed for our extensive experiments . . . . .	181
6.2	Hyperparameter $\alpha$ used for each approach based on $SM S_{EXP3}^\alpha$ and each level of certainty $\eta_v = \eta_b$ in a SAA-inc game ( $n = 3, m = 9, \varepsilon = 1$ ) . . . . .	182
A.1	<i>Example 1</i> . . . . .	221
C.1	<i>Example 1</i> . . . . .	295



# List of Algorithms

1	UCB	77
2	EXP3	79
3	Hash function for d-SAA	112
4	Example of function $h_1$	137
5	$DSMS_{EXP3}^\alpha$	166
6	$SDSMS_{EXP3}^\alpha$	171



# Introduction

For several decades, mobile operators have been battling for cellular spectrum in order to improve their quality of service and develop better wireless communication networks. In fact, much like a painter needs several colours to create their finest works, a mobile operator requires a broad spectrum of frequencies to provide the best possible service quality. For instance, low-frequency waves can transmit information over longer distances and penetrate dense obstacles more effectively than high-frequency waves. However, they provide lower data rates. Hence, to ensure high quality service, it is essential for an operator to have access to both frequency bands.

Nowadays, most countries decide to allocate their spectrum in the form of licences through auctions. For example, 5G licences were sold recently through auctions in many countries such as Austria, Sweden, United Kingdom, Germany, Norway, Denmark, France and the United States [[Dr Gerdis Marquardt, 2021](#); [Arcep, 2020](#); [FCC, 2020](#)]. Over the last 20 years, Simultaneous Ascending Auction (SAA), also known as Simultaneous Multi Round Auctions (SMRA), has become the privileged mechanism used for the sale of licensed cellular spectrum. This auction has the particularity of having a dynamic multi-round format in which bids are submitted simultaneously on all licences. It ends when no new bids have been submitted during a round. Its popularity is due to its relatively simple rules as well as the generation of substantial revenue for the regulator. It is believed to play a central role in the allocation of 6G frequencies which are expected to take place around 2030 [[Ministère, 2023](#)].

Considering the fact that sometimes billions of euros are at stake in SAA and that an operator's business plan highly relies on the acquired licences, it is crucial for an operator to have an efficient bidding strategy. However, auction theory or exact game resolution methods are unable to compute the optimal bidding strategy due to the high complexity of the induced bidding game [[Reeves, 2005](#)]. Moreover, a number of strategical issues add up

due to the mechanism of SAA and the specific condition of the operators such as budget constraints.

In this thesis, we propose an efficient bidding algorithm based on *Monte Carlo Tree Search* (MCTS) [Browne, 2012] and on the prediction of final auction prices to tackle simultaneously the main strategical issues of SAA. In order to construct such an algorithm, we decide to proceed step by step by gradually increasing the complexity of our SAA model throughout the chapters. For instance, various constraints are added progressively.

This thesis is organised as follows:

- Chapter 1 provides an overview of various auction mechanisms that have been employed over the years to sell cellular spectrum, highlighting their respective advantages and drawbacks.
- Chapter 2 presents the bidding problem in SAA, the relevant related work aimed at trying to solve it and the positioning of this thesis.
- Chapter 3 offers a concise survey of adversarial search methods, ranging from the minimax search algorithm to MCTS.
- In Chapter 4, we consider a turn-based deterministic model of SAA with no hidden information. We introduce an efficient bidding algorithm based on MCTS and a novel method to predicting final auction prices.
- In Chapter 5, we extend this model to simultaneous bidding, budget constraints, and other specificities pertinent to SAA. Bidders' characteristics, such as budgets, are considered common knowledge. In this context, we propose an efficient bidding algorithm based on MCTS and an adaptation of our price prediction method discussed in the previous chapter. Realistic-sized instances are used for experimental evaluation.
- In Chapter 6, we consider the scenario where bidders possess private information, making budgets no longer common knowledge. Here, we present three distinct approaches that adapt the bidding algorithm from the previous chapter.

# An introduction to spectrum auctions

## Contents

---

1.1	Assigning spectrum licences . . . . .	11
1.2	Spectrum auctions: objectives, properties, issues . . . . .	12
1.2.1	Objectives . . . . .	13
1.2.2	Properties . . . . .	13
1.2.3	Main strategical problems . . . . .	14
1.3	Main families of spectrum auctions . . . . .	16
1.3.1	Single round sealed-bid auctions . . . . .	16
1.3.2	Multi-round auctions . . . . .	20

---

Handing cellular spectrum to companies is divided into two steps. The first step consists in associating the spectrum to licenses. Each licence is defined by its frequency band, its geographic coverage, its period of usage and its restrictions on use. The second step consists in assigning the licences to the different companies.

## 1.1 Assigning spectrum licences

According to [Cramton, 2002], there are three main assignation methods:

- The historical approach is often referred to as a **beauty contest** as each participant submits a proposal on how it plans to use the spectrum and, then, the regulator assigns the spectrum licences accordingly to the participants with the most attractive proposals. However, this process has two downsides. First, the process is extremely slow. Secondly, as the assignation process is inherently subjective, it can easily be subject to corruption and, thus, makes it contestable. Nevertheless, it is still used for the assignation of 5G licences in many countries such as Japan [European 5G Observatory, 2019] or Singapore [Infocomm, 2020].



- An alternative is to randomly select licence winners from amongst those that apply. This approach, known as **lotteries**, seems to be fairer and faster than beauty contests. Hence, this is the main reason why the Federal Communication Commission (FCC) in the United States decided to switch from beauty contests to lotteries in the 1980s. However, the lotteries, held between 1986 and 1989, offering 643 cellular licences attracted over 320 000 applicants as they were considered extremely valuable and any U.S citizen could participate [Hazlett, 1993]. In addition to the calamitous processing of the high number of applications, the winners were not the best suited for exploiting the spectrum. It took years for the licences to be transferred to those capable of using them, often through private auctions. Lotteries were then quickly abandoned by the FCC.
- A third approach is to rely on **auctions**. These present many advantages. First, all participants are well aware of the award criteria which makes them fair. As they are open and transparent, they can not raise suspicion and be subject to allegation of corruption. Secondly, the winners are supposed to be the best suited to develop the network. Indeed, companies with ambitious projects and high values for the licences should bid higher than other participants and, hence, win the licences. To ensure that the assigned spectrum is correctly and equally deployed on the territory, regulatory authorities can ask winners to respect a certain number of obligations. For example, in France, the *Autorité de Régulation des Communications Électroniques, des Postes et de la Distribution de la Presse* (ARCEP) verifies that each 5G licence winner deploys a minimum number of sites each year. Moreover, operators are constrained to have an homogeneous deployment across the country, including in rural areas. Hence, this complementary process is supposed to deter any ill-intentioned bidders. Finally, competition is no longer a burden for the assignation of licences. On the contrary, a higher competition leads to higher bids and, hence, to higher auction revenues. Nowadays, spectrum licences are mainly assigned through auctions.

## 1.2 Spectrum auctions: objectives, properties, issues

The first question which comes to mind when selling spectrum licences is to who and at what price. In 1959, Ronald Coase proposed to assign them through auctions [Coase, 1959]. However, it was only thirty years after that the FCC decided to hold its first spectrum auction in the United States. In this section, we start by stating the two main objectives of an auction.

Then, we present some interesting properties induced by an auction mechanism. Finally, we shed light on the main strategical problems that a bidder can face in an auction.

### 1.2.1 Objectives

Auctions have two objectives. The main one is to assign the licences to those best suited to provide service and develop the network. This is often referred to as *efficiency*. The secondary objective is to maximise the regulator's revenue. However, this revenue should not be maximised to the point where it prevents companies to develop their network due to too high acquisition costs. Moreover, a spectrum cap, i.e. a limit of spectrum that an entity can hold, is often imposed to bidders to avoid monopoly and encourage competition for the development of the network. This generally reduces revenue.

To achieve those objectives, the respective authorities need to distribute the spectrum intelligently between licences and then adopt an efficient auction mechanism for their assignment.

### 1.2.2 Properties

An auction can be designed to satisfy some theoretical properties facilitating the achievement of both objectives, efficiency and revenue maximisation. We present some properties here below [[Nedelec, 2022](#)]. A mechanism is:

- **Bayesian Incentive-compatible:** If the profile of strategies where all bidders bid truthfully, i.e. bid their true value, is a Bayesian Nash equilibrium [[Myerson, 1978](#)]. This means that if a bidder's opponents all bid truthfully, then this bidder's best response, i.e. the bidding strategy which maximises its utility, is to bid truthfully.
- **Truthful or Dominant Strategy Incentive-Compatible:** If bidding truthfully is a weakly dominant strategy for all bidders. This means that, regardless of what the other bidders bid, bidding truthfully guarantees a bidder a payoff at least as high as any other bidding strategy. Hence, a truthful mechanism is a Bayesian Incentive-compatible mechanism. The converse is not true.
- **Standard:** If each licence is assigned to the bidder which has submitted the highest bid.

- **Efficient:** If each licence is assigned to the bidder which values it the most.
- **Individually-Rational:** If a bidder weakly prefers participation in an auction to not participating, i.e. the bidder should never overpay for what they win (assuming they bid truthfully) and should, hence, obtain a positive utility (at least in expectation in the case of incomplete information).
- **Fair:** If all bidders have a similar probability of winning regardless of their wealth. This egalitarian social welfare approach is mainly used in recurrent auctions to reduce the bidder drop problem and maximise revenue in the long run. For instance, if licences are always allocated to the same bidders, then the rest of the bidders become unsatisfied and may decide to leave the market. Hence, only the most powerful bidders remain. There is a high risk that they will decide to create an oligopoly, provoking a fall in prices that can collapse the market and bankrupt the auctioneer. Moreover, even if the auctioneer decides to set reservation prices, many licences might not be sold reducing the efficiency and the revenue of the auctioneer. Maintaining a certain level of competition also protects the consumers. Indeed, in the specific case of an oligopoly, companies could all fix high prices and the consumer would have no other choice but to be ripped off. Many measures of fairness can be used by the auctioneer for allocating licences such as Jain's index [Jain, 1984].

These properties have different effects on the auction's outcome and, hence, on its objectives. For example, by using a truthful mechanism, bidders have an incentive to bid their true value. Given this, it might be easier to maximise efficiency and revenue.

### 1.2.3 Main strategical problems

Each mechanism comes with its pros and cons. In this section, we present in detail the main strategical problems that a mechanism can generate.

- **Winner's curse:** It is the fact of overpaying a licence in a competitive auction, i.e. paying more than its intrinsic value. Hence, the winner ends up with a negative utility. This is generally due to incomplete information [Bazerman, 1983] where the winner is the bidder who has overestimated the most the value of a licence. Other reasons such as pressure or emotions can also cause the winner's curse [Foreman, 1996]. The winner's curse sometimes refers to the disappointment of a winner as its licence is worth less than expected [Thaler, 1988]. The major difference with the first definition

is that, according to the second, the winner can still make profit. The profit will just be less than expected.

- **Exposure problem:** This issue occurs when a bidder pursues a set of complementary licences, i.e. a collection of licences which is worth more than the sum of the licences individually. By bidding more than its individual value on each licence, a bidder faces the risk of paying too much for the subset of licences it actually wins if the rest of licences are won by other bidders. Hence, an exposed bidder obtains a negative utility. Many strategic reasons can explain that some licences are worth more together than individually. For example, a licence for a low-frequency band and a licence for a high-frequency band are usually considered as complements. This is due to the intrinsic characteristics of each radio wave. Low-frequency waves can transmit information over greater distances and pass through dense obstacles more easily than high frequency waves. Hence, they are ideal to cover lightly populated zones in rural environments. However, low-frequencies generally offer smaller data rates. Hence, to improve quality of service and offer a high-speed connection to its users, an operator needs to also rely on high-frequency waves. In order to be competitive, an operator needs both radio waves today. This problem can be eliminated by using a combinatorial auction which offers the possibility to bidders to bid directly on a combination of licences rather than bidding on each licence individually.
- **Threshold problem:** This issue is introduced by the use of a combinatorial auction [Rothkopf, 1998] and refers to a variant of the free-rider problem [Milgrom, 2000]. More precisely, it corresponds to the situation where an individual bidder is unable to displace an inefficient package bid without the help of other bidders. This problem is often due to the fact that bidders which bid on small combination of licences are unaware of each other's presence. Hence, given this lack of information, they are incapable of outbidding a bidder bidding on a large set of licences, even if their valuation allows them to do so. Most of the literature discusses the impact of the threshold problem in multi-round combinatorial auctions [Pekeč, 2003; Bykowski, 2000; Cramton, 2002]. However, its impact is still relevant in single round sealed-bids combinatorial auctions [Vangerven, 2021].

*Example.* Suppose a multi-round combinatorial auction with 2 items, A and B, between 3 bidders. No bidder values any asset except as follows: licence A is worth 50\$ for bidder 1, licence B is worth 50\$ for bidder 2 and the combination of both licences is worth 90\$ for bidder 3. It is important to note that these valuations are private information. The efficient allocation is to give licence A to bidder 1 and licence B to

bidder 2. During the first round, suppose bidder 1 bids 10\$ on  $A$ , bidder 2 bids 10\$ on  $B$  and bidder 3 bids 70\$ on both licences. Both licences are then temporarily won by bidder 3. Moreover, neither bidder 1 nor bidder 2 can single handedly overcome the difference of 50\$ between the winning bid and the sum of their bids. This single handedly insurmountable difference is named the *threshold*. As none of the bidders are aware of their opponents' valuation, there is a fair chance that bidder 1 and bidder 2 are unable to cooperate to overcome the threshold and that bidder 3 ends up by winning both licences.

The threshold problem and exposure problem are often considered as opposing forces [Rothkopf, 1998]. The first favours bidders bidding on large combination licences while the second favours bidder bidding on small combination of licences. However, depending on their severity, they can both have a big impact on an auction's efficiency as well as on its revenue. When synergies between licences are important, it is commonly admitted that the exposure problem has a bigger impact on the auction's objectives than the threshold problem. The winner's curse can also impact an auction's outcome as bidders may shade their bids to avoid it.

## 1.3 Main families of spectrum auctions

In this section, we present the main families of spectrum auction mechanisms which have been used until today. We start by presenting mechanisms for single round sealed-bid auctions and, then, we will present mechanisms for multi-round auctions.

### 1.3.1 Single round sealed-bid auctions

Single round sealed-bid auctions consist of one single bidding round where all bidders submit their bids simultaneously. This puts a lot of pressure on bidders as there is no comeback in case of poor bidding. It is impossible for them to adjust their bids based on competing bids as it is the case in multi-round auctions. Hence, as a bidder has no information on its opponents' bids, it is essentially bidding blindly. However, single round auctions have the advantage of being an extremely fast assignment process compared to some multi-round auctions which can drag on for months. For instance, the recent Portuguese 5G multi-round auction concluded after 200 days and 1727 bidding rounds [ANACOM, 2021]. There are mainly two types of single round auctions used for the

assignment of spectrum: first-price auctions and second-price auctions. They differ from one another by their pricing function.

#### 1.3.1.1 First price sealed-bid auction

In a single-item first price sealed-bid auction, the highest bidder wins the licence and pays its actual bid. The fact that the winner pays literally its bid and can not adjust it subsequently can lead to undesired consequences. Indeed, this lack of information can lead to overbidding, i.e. paying substantially more than what was necessary to acquire a specific licence. For instance, in the 2G Brazilian spectrum auction, BellSouth ended up paying 1 billion dollars more than the next highest bidder for the Sao Paulo Metro licence [GSMA, 2014]. To avoid this issue, bidders have a tendency to shade their bids at the risk of not winning the licence. Hence, it is highly possible that the winner is not the most suited for the licence.

#### 1.3.1.2 Sealed-bid second-price auction

In a single-item sealed-bid second-price auction, also known as Vickrey auction [Vickrey, 1961], the highest bidder wins the licence and pays the second highest bid. The fact that the winner does not pay its bid but the second highest bid encourages each bidder to bid its true value. Indeed, in a Vickrey auction, truthful bidding, i.e. bidding its true value, is a weakly dominant strategy. Hence, the licence should be assigned to the bidder with the highest value. In practice, bidders do not necessarily bid their true value. Firstly, this can be due to the fact that bidders might have uncertainties about the actual value of a licence's worth. Depending on their strategy, they might decide to overbid or underbid according to this uncertainty. Secondly, as pointed out in [Rothkopf, 1990], bidders might fear that their true value will be revealed to third parties with which they interact after the auction. Indeed, such information could indicate to other firms how much a bidder is ready to spend and, hence, how much they can yield in future negotiations. Thirdly, bidders might also fear that the auctioneer might cheat by pretending that another bid was received just under the highest bid [Lucking-Reiley, 2000]. This last issue might not be so important in spectrum auctions where we imagine the general process is relatively controlled. However, this is an issue which is intrinsic to Vickrey auctions and can not be excluded in the general case. Moreover, regarding the regulator, using this auction mechanism is not necessarily a good choice, especially when there is a huge difference between the values of the bidders. Indeed, the winner can end up paying only a small fraction of what it was willing to pay for the licence. For instance, the sealed-bid second-price spectrum auction was the cause

of national political embarrassment in New Zealand in 1990 as bidders paid far less than their bids [McMillan, 1994]. In one of the auctions, the winner bid 100 000\$NZ and got the licence for 6\$NZ. In another, a company paid 5000\$NZ for a licence and had bid 7 million \$NZ. The National Economic Research Associates, which advised the New Zealand government to use the Vickrey auction, had predicted initially a revenue of 240 million \$NZ but it only generated 36 million \$NZ.

A fundamental question which might be raised is "Which of these two auctions generates the highest revenue?". If the bidders are risk-neutral, have independent-private-values, have common beliefs about the other valuations and their payments is a function of their bids alone, then the first price sealed-bid auction and Vickrey auction generate the same expected revenue [Vickrey, 1961]. This is known as the *revenue equivalence theorem*. However, these conditions are rarely met in practise. Indeed, bidders are generally risk-averse as they prefer generally lower utilities if they are more certain. In this specific case, it has been shown that the first price sealed-bid auction generates higher revenue than the Vickrey auction [Waehrer, 1998].

### 1.3.1.3 Multi-item auctions

Spectrum auctions usually involve more than one licence. They are two ways of extending the above single round auctions to multiple licences. Either by running one independent auction for each licence or by running a combinatorial auction for a set of licences.

#### A) Running multiple independent auctions

Running  $m$  independent auctions for  $m$  licences has the following issues. First, identical licences can be sold at very different prices. For instance, in the Vickrey auction held in New Zealand in 1990, Totalisator Agency Board bid 401 000\$NZ and obtained a licence for 100 000\$NZ while BCL bid 255 124\$NZ and had to pay 200 000\$NZ for an identical licence. Hence, the company which wanted to pay more was charged two times less for the same licence [Milgrom, 2004]. A second issue is the *exposure problem* which has already been presented in Section 1.2.3.

## B) Combinatorial auctions

Combinatorial auctions, often referred to as package bidding, offer the possibility to bidders to bid directly on a combination of licences rather than bidding on each licence individually. By bidding on a combination of licences, the bidder either wins the whole combination or nothing. Hence, combinatorial auctions do not suffer from the exposure problem. However, they create two new issues.

The first issue is the *threshold problem*, presented in Section 1.2.3, which can heavily impact an auction's efficiency and revenue.

The second issue introduced by combinatorial auctions concerns the *complexity* of the winner determination problem: "Given a set of bids in a combinatorial auction, find an allocation of items to bidders (the auctioneer can keep some of the items) that maximises the auctioneer's revenue" [Cramton, 2007]. As this problem is NP-complete, identifying the revenue maximising assignment becomes computationally intractable for large instances. One proposed solution is to restrict the set of combinations on which bidders can bid [Rothkopf, 1998]. However, this may eliminate desirable combinations and reintroduces the exposure problem.

According to the report published by GSMA in 2014 [GSMA, 2014], combinatorial first-price auctions are the most distorted of the main auction approaches and should be avoided. Indeed, they present a certain number of inefficiencies such as the winner's curse [Thaler, 1988], increased knock-out risks or price disparities. Nevertheless, they were still used a decade ago for the allocation of spectrum licences such as 3G in France (2011) or 4G in Norway (2013) [Kokott, 2017].

## C) Vickrey-Clarke-Groves auction

The most famous combinatorial auction is the Vickrey-Clarke- Groves (VCG) auction [Lucking-Reiley, 2000]. It is a generalisation of the second-price auction. As for any combinatorial auction, the optimal allocation is computed through the winner determination problem. Its payment rule relies on the notion of social welfare. Theoretically, social welfare is defined as the sum of the seller revenue and the bidder payoffs for a given allocation. However, in VCG, it is commonly used to design the sum of bids of a given allocation.



Hence, the optimal social welfare corresponds to the sum of bids obtained for the optimal allocation. At the end of the auction, each bidder must pay its social cost, i.e. the loss in social welfare incurred by the rest of the bidders due to its presence in the auction. With this specific payment rule, VCG induces a truthful mechanism. In other words, as it was the case for the single-item Vickrey auction, bidders are encouraged to bid their true valuation.

However, this mechanism has many downsides. A first downside is that determining the optimal allocation is NP-hard in general. A second major downside is that the auctioneer revenue may be zero even if all bidders place high bids.

*Example.* Consider a VCG auction with 4 bidders and 3 licences  $l_1$ ,  $l_2$  and  $l_3$ . No bidder values any licence except as follows:  $\{l_1\}$  is worth 100\$ for bidder 1,  $\{l_2\}$  is worth 100\$ for bidder 2,  $\{l_3\}$  is worth 100\$ for bidder 3 and  $\{l_1, l_2, l_3\}$  is worth 200\$ for bidder 4. If all bidders play their true value, bidder 1 wins  $l_1$ , bidder 2 wins  $l_2$  and bidder 3 wins  $l_3$ . However, each bidder pays zero.

In [Rothkopf, 2007], Rothkopf lists 13 other reasons why VCG does not work well in practise. Nevertheless, it is still used today for the assignment of 5G licences, generally in the case of small instances with relatively cheap licences. For instance, this mechanism is used in many French overseas for the assignment of 5G licences such as La Reunion or Mayotte in 2022 [Arcep, 2021].

### 1.3.2 Multi-round auctions

The main advantage of multi-round auctions is that the bidding process reveals information about each bidder's valuation. Hence, by taking it into account and updating their current likelihood of acquiring different combination of licences, bidders can progressively adapt their bids. In some situations, this can stimulate competition as it reduces the winner's curse, hence bidders can safely bid more aggressively. However, this information can also be used to facilitate collusions. Indeed, bidders have generally a common interest in keeping prices low and, therefore, they might take advantage of the bidding process to implicitly agree on a split of licences. Moreover, it is possible to enforce these tacit agreements by immediately punishing a bidder that has decided to deviate from the initial agreement. Such threats are not possible in single-round auctions and, hence, collusion is reduced. In addition to reducing the auctioneer's revenue, splitting up the licences through tacit agreements can also greatly impact an auction's efficiency as not all licences may be

acquired by those best suited to fully use them. Nevertheless, it is commonly accepted that the advantage of revealing more information in the bidding process outweighs the risk of collusion [Cramton, 2002]. Today, multi-round spectrum auctions generally use one of the following formats or variations thereof: Simultaneous Ascending Auction (SAA) or Combinatorial Clock Auction (CCA). We present in detail both of these mechanisms hereafter.

### 1.3.2.1 Simultaneous Ascending Auction

#### A) Mechanism

Simultaneous Ascending Auction (SAA) [Milgrom, 2000; Cramton, 2006], also known as Simultaneous Multi Round Auction, is a dynamic multi-round auction mechanism where each licence is sold via a separate and concurrent English auction. At each round, bidders simultaneously submit their bids for each licence they are interested in. The bid price of licence  $j$ , denoted  $P_j$ , corresponds to the highest bid obtained so far for licence  $j$  and, thus, is its current selling price. At the beginning of each round, the bid price and the current winner of each licence is announced to each bidder. A new bid  $P_j^{new}$  for licence  $j$  is only accepted if it is higher than the current bid price  $P_j$  plus a small fixed amount  $\varepsilon$  named bid increment:  $P_j^{new} \geq P_j + \varepsilon$ . Submitting a bid lower than this amount is equivalent to not bidding. At the end of each round, the temporary winner of each licence is designated as the bidder who submitted the highest admissible bid and the bid price is updated. As all bids are submitted simultaneously, ties can occur between bidders having bid the same amount on a given licence. The usual tie-breaking rule is to select randomly the temporary winner amongst these bidders. The auction ends when no new admissible bids have been submitted on any licence during a round [Milgrom, 2000]. Each licence is then sold at its bid price to its corresponding winner.

#### B) Activity rules

Activity rules are introduced in SAA to penalise bidders which don't maintain a certain level of bidding activity. At the beginning of the auction, each bidder is given a certain level of eligibility. This level determines on which licences a bidder is allowed to bid on. The minimum level of eligibility required to bid on a given licence is tightly correlated to the estimation of its value by the auctioneer [Milgrom, 2000]. Each round a bidder fails

to match the corresponding activity rules, its eligibility is reduced. Hence, during the course of an auction, there are sets of licences which a bidder can no longer bid on as they have exceeded its current eligibility. These activity rules have mainly two functions [Milgrom, 2000]. First, they put pressure on bidders to bid actively and, hence, fasten up the pace of the auction. They limit wait-and-see strategies where bidders wait before placing any serious bids in order to get finer estimates of one's valuations and take advantage of competitors who have already committed an important part of their budget in acquiring other licences. Indeed, if a bidder remains too inactive, it might become ineligible to bid in later rounds. Therefore, activity rules push bidders to reveal their true intentions early on in the auction. Secondly, by communicating each one's eligibility at the beginning of each round, it increases the amount of information available to bidders to develop their strategy. As long as there is no collusions, more information should mean higher efficiency and revenue [Cramton, 2002]. Each bidder is given five "waivers" of the activity rule [Milgrom, 2000] which, when used, avoids a reduction in its eligibility in a given round. They were initially introduced in case of a bidder error. However, they can also be used strategically.

### C) Popularity

The SAA was first introduced in 1994 by the US Federal Communications Commission (FCC) for the allocation of wireless spectrum rights and generated a revenue of \$617 million. Since this enormous success, it has become the most privileged mechanism used for spectrum auctions. For instance, it has been used in many countries for the assignment of 5G licences such as Portugal [ANACOM, 2021], Italy [European 5G Observatory, 2018] or Germany [Bundesnetzagentur, 2022]. It is also used today as the principal stage of a two-step spectrum auction. The principal stage consists in determining the amount of spectrum won by each bidder in each band and the assignment stage step consists in determining the precise frequencies of lots won in the principal stage. For example, SAA was used as the principal stage in the 5G UK auction [Ofcom, 2020]. It was initially designed by Paul Milgrom and Robert Wilson who both received the 2020 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, mainly for their contributions to the SAA.

## D) Pros and cons

Here a few reasons which explains SAA popularity in spectrum auctions. First, revenue in past auctions have been substantial. For instance, revenue in past US auctions have exceeded the government estimates [Cramton, 2006]. These large revenues can be partially explained by the open bidding and multi-round nature of SAA. Indeed, bidders can freely adjust their bids throughout the auction while taking into account the latest information about the likelihood of acquiring different combination of licences. Hence, bidders can progressively adapt their strategies during the whole course of the auction. This reduces the winner's curse and encourages bidders to bid more aggressively. Secondly, identical licences are sold for similar prices. Thirdly, it appears that bidders successfully acquire efficient combinations of licences [Cramton, 2006], i.e. sets of complementary licences. This is notably due to the two main features of SAA: (1) the use of multiple rounds and (2) simultaneous sales, rather than sequential. Finally, a very appealing characteristic of SAA for auctioneers is the simplicity of its rules.

However, as SAA is not a combinatorial auction, it suffers from the exposure problem. This can impact the efficiency and revenue of an auction as risk-averse bidders may prefer bidding safe than incurring the risk of being exposed in pursuing complementary licences. Moreover, efficiency and revenue are also impacted by demand reduction and collusive bidding. Indeed, it is sometimes more profitable for a bidder to reduce its demand to guarantee buying licences at a low price. These tactics to avoid price rise are facilitated by the open bidding process where public information can be used to communicate, especially in early rounds. The most clear example is surely the second-generation GSM spectrum auction which took place in Germany in 1999 [Grimm, 2003]. In this auction, ten licences were sold (9 identical ones with bandwidth  $2 \times 1$  and a larger one with bandwidth  $2 \times 1.4$  MHz) between four incumbent operators. In the first round, Mannesmann performed surprisingly high jump bids by bidding DM 36.36 million on the first five smaller licences, DM 40 million on the four next smaller licences and DM 56 million on the larger frequency band. This is a typical example of how jump bidding can be used for signalling. This suggests to the other stronger incumbent, T-Mobile, to only bid on the first five licences so they can each buy the small licences at DM 40 million. To ensure that T-Mobile does not feel ripped off for not obtaining the larger frequency band, Mannesmann bided DM 56 million on it which is the same price paid by T-Mobile by MHz for the five smaller licences. In the second round, T-Mobile bided slightly more than the bid increment on the five first licences. The two other incumbents decided to drop out. In the third round, no bids were submitted and the auction

closed. The bid price of each licence as well as their corresponding temporary winner for each bidding round are summed up in Table 1.1.

Licences	1	2	3	4	5	6	7	8	9	10
	R1									
Bid price	36.36	36.36	36.36	36.36	36.36	40	40	40	40	56
Winner	M	M	M	M	M	M	M	M	M	M
	R2									
Bid price	40.01	40.01	40.01	40.01	40.01	40	40	40	40	56
Winner	T	T	T	T	T	M	M	M	M	M
	R3									
Bid price	40.01	40.01	40.01	40.01	40.01	40	40	40	40	56
Winner	T	T	T	T	T	M	M	M	M	M

**Tab. 1.1.:** The GSM spectrum auction in Germany in 1999. Licences 1 to 9 correspond to bandwidth  $2 \times 1$  and licence 10 to  $2 \times 1.4$  MHz. Winners are either Mannesmann (M) or T-Mobile (T).

Bidders can also exploit the trailing digits to coordinate themselves and specify which licences they want. This is known as code bidding. As bids are usually in millions of dollars, using the last three digits is considered as a minimal cost to communicate [Cramton, 2001]. For instance, in the GSM Spectrum Auction in Germany in 1999, the bid of DM 36.36 million of the five first smaller licences with the repetition of number 36 can easily be interpreted as a proposal of splitting up the licences into two.

Additional rules are sometimes imposed in SAA to reduce the above issues. For instance, click-box bidding is often used in SAA nowadays. At each round, bidders no longer submit whatever value of bid they like but have to indicate in a click-box the number of bid increments (often from 1-9 [Cramton, 2002]) they wish to bid above the current bid price of a licence. This prevents code bidding. Moreover, by reducing the maximum number of bid increments than one can place above a licence's bid price, signalling through jump bidding can also be eliminated. However, click-box bidding increases the number of tie bids. It is generally considered by the auctioneers as a smaller issue as, even if ties happen relatively often at the beginning of the auction, it is unlikely that the final bid on a licence involves a tie.

### 1.3.2.2 Combinatorial Clock Auction

#### A) Mechanism

Combinatorial Clock Auction (CCA) is a two-stage dynamic bidding process which was first proposed by Ausubel *et al.* [Ausubel, 2006a] in 2006. The first stage is known as the *clock rounds* and the second stage as the *supplementary round*.

- *Clock rounds*: It is a multiple round clock auction. At each round, the bid prices are announced and bidders state their demands for each licence. In other words, each bidder submits a unique package bid each round. If the sum of demands of all bidders exceeds supply for a given licence, then the auctioneer increases its bid price. Otherwise, the bid price stays the same. This bidding process ends when it reaches a round with no excess demand.
- *Supplementary round*: It corresponds to the auction final round. After the clock rounds have ended, bidders are asked to participate in a single-round combinatorial auction to submit additional package bids. In order to be admissible, these package bids need to be greater than their corresponding price obtained at the end of the clock rounds.

Throughout the two stages, all bids are treated as all-or-nothing package bids. As it was the case for single-round combinatorial auctions, the licences are assigned according to the allocation which maximises the auctioneer's revenue. This allocation is computed through the winner determination problem by taking all package bids submitted during the clock rounds and the supplementary round. Hence, it is possible that the supplementary round has no impact on the assignment of licences. Moreover, the package bids of a specific bidder are mutually exclusive, i.e. only one package bid of a specific bidder can be part of the revenue maximising assignment. Depending on the final allocation, each bidder needs to pay a certain amount. Different payment rules have been proposed in the literature and are discussed in the next section.

Today, an extra stage is often added to the two initial stages of CCA [Ausubel, 2017]. In this case, licences are grouped in categories of equal values. In the clock rounds and supplementary round, bidders compete for a number of licences in each category. The third stage is known as *assignment stage*.

- *Assignment stage:* This stage consists in mapping the number of generic licences obtained in each category during the two first stages to specific physical frequencies. To do so, a single-round sealed-bid auction is usually performed.

## B) Payment rule

Surprisingly, the choice of payment rule to be used after the two first stages of CCA is still being discussed today. One solution would be to apply VCG prices as it is done for second-price sealed-bid auctions. By doing so, it should create an incentive for bidders to bid their true values, hence maximising the auction's efficiency. However, as it has already been discussed, this payment rule can generate low revenue for the auctioneer even with high competition. More precisely, when licences exhibit complementarities, VCG often produces a price vector outside of the core [Ausubel, 2006b], i.e. there exists a coalition of bidders willing to pay more than the total price paid by the current winners to the auctioneer. For this reason, CCA generally relies on a core-selecting payment rule.

Many different core-selecting payment rules have been proposed [Bünz, 2022]. The one which is most used in practise for spectrum auctions and has generated over \$20 billion revenue is known as the Quadratic rule [Day, 2012]. The final price determination works as follows: (1) Consider only price vectors which are in the core (2) Among these price vectors, consider only the ones which minimises the auctioneer's revenue (3) Finally, among the remaining price vectors, select one which minimises the Euclidean distance with the VCG price vector.

## C) Objectives

The clock phase has as main objective to produce highly useable price discovery. At each round, a bidder gets a finer estimate of the final price of each licence and can then focus on bidding on the package which maximises its valuation. This improves the auction's efficiency. Moreover, bidders have only access to aggregate information and not to individual bids. Hence, it is impossible to communicate with one another using bid signalling. Collusive strategies are also eliminated as one can not know if a bidder has deviated or not from the initial agreement. However, in the last rounds of the clock phase, bidders may benefit from demand reduction to avoid price rise. This reduces revenue and efficiency. The role of

the supplementary stage is to eliminate this inefficiency by inciting bidders to bid their true value. Moreover, as the last round is a single-round sealed-bid auction, each bidder has interest in betraying their collusion as no punishment can then be inflicted. Finally, CCA completely eliminates the exposure problem as package bids are used in all stages.

## D) Activity rules

Without any additional rules, bidders may have interest in underbidding or sabotaging the clock phase. Indeed, by concealing their true interests from their opponents, they can easily submit surprise bids in the supplementary round. This is often referred as bid snipping. Moreover, a bidder may have interest in sabotaging the clock phase and increasing its opponents' payments by overbidding on licences it does not want to have. To avoid these issues and ensure sincere bidding, activity rules are needed in two places of the CCA [Ausubel, 2017]:

- *Clock round activity rule:* To ensure consistent bidding through out the clock phase. Bidding in later rounds need to be coherent with previous bids.
- *Supplementary round activity rule:* To prevent bidders from deviating drastically from their clock phase bids and, hence, ensure sincere bidding during the clock phase.

Eligibility-Point Monotonicity is often used as a clock round activity rule. It follows the same principle as the activity rule described in SAA with different eligibility points attributed to each licence. However, this activity rule has a few downsides. One of them is that bidders may bid on undesired cheap licences just to maintain their eligibility level. This is known as "parking" [Ausubel, 2006a] and compromises price discovery. Another issue is that it can prevent a bidder from pursuing the package that maximises its utility whenever this package exceeds its eligibility, even though this bidder has always bid sincerely [Ausubel, 2011]. For instance, suppose licences A and B are substitutes for a given bidder and that licence B requires more eligibility points than A. If this specific bidder decides to pursue A at the beginning of the auction, its eligibility is reduced. However, if the bid price of licence A skyrockets, then this bidder is unable to pursue licence B as it now exceeds its eligibility level. To avoid such issues and ensure coherent sincere bidding, the following clock round activity rule was proposed in [Ausubel, 2006a].



**Definition 1.1.** *Revealed preference activity rule (RP):* Suppose two different rounds,  $s$  and  $t$ , with  $s < t$ . Let  $p^s$  and  $p^t$  be the price vector at these rounds. Let  $x^s$  and  $x^t$  be the associate demands of a given bidder. Let  $v(x)$  be the value given to package  $x$  by this specific bidder. For a sincere bidder, preferring  $x^s$  to  $x^t$  when prices are  $p^s$  means:

$$v(x^s) - p^s \cdot x^s \geq v(x^t) - p^s \cdot x^t$$

and preferring  $x^t$  to  $x^s$  when prices are  $p^t$  means:

$$v(x^t) - p^t \cdot x^t \geq v(x^s) - p^t \cdot x^s$$

By summing up these inequalities, if a bidder bids sincerely then:

$$(p^t - p^s) \cdot (x^t - x^s) \leq 0$$

The above inequality corresponds to the revealed preference activity rule and, to ensure a bidder's sincerity, a bidder's demand  $x^t$  must satisfy this inequality for all rounds  $s < t$ .

In practise, the revealed preference activity rule was rarely used in the clock phase. However, recent approaches tend to take in account eligibility as well as revealed preferences [Ausubel, 2017].

Concerning the supplementary round activity rule, Ausubel *et al.* [Ausubel, 2006a] initially proposed a relaxed revealed-preference activity rule (RRP):

$$\alpha(p^t \cdot x^s - p^s \cdot x^s) \leq p^t \cdot x^t - p^s \cdot x^t \quad \text{with } s < t \text{ and } \alpha > 1$$

This inequality is applied to every round  $s$  in the clock phase and round  $t$  corresponds to the supplementary round. Using directly (RP) as the supplementary round activity rule would have been too strict and would have prevented bidders from bidding on the packages which maximise their utility. Indeed, to avoid the bid prices to rise too much, bidders often rely on demand reduction. However, the aim of the supplementary round is to eliminate this inefficiency and, hence, it is important to allow bidders to expand their demands to a certain extent. In practise, the supplementary round activity rule used in spectrum auctions is often a combination of (RP) and eligibility points [Ausubel, 2017]. The main difference between the below rules is the choice of the clock rounds used for the application of (RP).

- *Relative cap*: A bidder can only bid on a package  $x$  in the supplementary round if it satisfies (RP) with respect to the last clock round where it had bid on a package having at least the same level of eligibility as  $x$ .
- *Intermediate cap*: A bidder can only bid on package  $x$  in the supplementary round if it satisfies (RP) with respect to all rounds where it reduced its eligibility starting from the last clock round where it had bid on a package having at least the same level of eligibility as  $x$ .
- *Final cap*: A bidder can only bid on package  $x$  in the supplementary round if it satisfies (RP) with respect to the final clock round.

For instance, the supplementary round activity rule in the 4G UK Auction (2013) was relative cap, in Ireland Multi-Band Spectrum Auction (2012) was relative cap+final cap and in Canada 700 MHz Spectrum Auction (2014) was intermediate cap+final cap [Ausubel, 2017].

## E) Comparison with SAA

CCA has two main advantages over SAA. First, it considerably reduces collusive bidding and demand reduction. Hence, it is supposed to ensure a competitive outcome which increases the auction's efficiency and revenue. Secondly, it completely eliminates the exposure problem. Thus, CCA seems to have the potential of replacing SAA and becoming the new standard mechanism design for spectrum auctions. However, after having been massively used for the assignment of 4G licences, most countries have decided to switch back to SAA for the assignment of 5G licences [Dr Gerdis Marquardt, 2021]. For example, the UK, after having used CCA as its 4G spectrum auction, had initially announced a CCA for assigning its 5G licences in December 2018 but finally decided to hold an SAA. Why is this the case? This is mainly due to the fact that CCA is too complex. Even though it might lead to higher efficiency, complex mechanisms seem to be on the way out in favour of simplicity. Another important downside of CCA is the difference in price paid by companies for relatively similar spectrum. For instance, in the Swiss spectrum auction in 2012, Orange and Sunrise both bought 160 MHz but Sunrise paid 3 times more. Moreover, Sunrise paid 120 million Swiss francs more than Swisscom which had obtained more spectrum (225 MHz) [ComCom, 2012]. Hence, it seems that SAA will continue to be the most privileged mechanism used in spectrum auctions, at least in the near future. Nevertheless, some

countries such as Ireland [[ComReg, 2021](#)] or Israel [[Blumrosen, 2023](#)] still decided to use CCA for assigning their 5G licences.

# Bidding in Simultaneous Ascending Auctions

## Contents

---

2.1	Bidder's characteristics . . . . .	32
2.1.1	Utility and value function . . . . .	33
2.1.2	Budgets . . . . .	35
2.1.3	Eligibility . . . . .	36
2.2	Complexities induced by the bidding game in SAA . . . . .	36
2.2.1	General game properties . . . . .	37
2.2.2	Game complexities . . . . .	39
2.2.3	Strategical complexities . . . . .	40
2.3	Performance indicators . . . . .	44
2.4	Benchmarking: perceived-price bidding strategies . . . . .	46
2.4.1	Straightforward bidding . . . . .	47
2.4.2	Point-price prediction strategy . . . . .	49
2.4.3	Distribution price prediction strategy . . . . .	52
2.5	Other related works . . . . .	55
2.5.1	Exposure . . . . .	55
2.5.2	Own price effect . . . . .	57
2.5.3	Budget . . . . .	58
2.5.4	Eligibility management problem . . . . .	60
2.6	Positioning of the thesis . . . . .	60

---

In this thesis, we focus on how to bid efficiently in the Simultaneous Ascending Auction (SAA). This is an important issue for operators. The two main reasons are that:

- **Licences are expensive:** Purchasing licences through spectrum auction is a heavy investment for operators. Indeed, operators sometimes end up spending a few billions of euros to acquire licences, e.g. Deutsche Telekom spent 2.17 billion euros in the 5G German SAA [[Bundesnetzagentur, 2022](#)].

- **Acquired licences highly impact an operator’s business strategy:** For telecommunication operators, frequencies represent the sinews of war. Geographic coverage and data rates fully depend on the acquired licences. Therefore, an operator’s business strategy highly depends on the spectrum auctions’ outcomes. For example, in an extreme case, if an operator has no frequencies, then it cannot offer any coverage and, hence, will have no clients.

As presented in Section 1.3.2.1, SAA has a dynamic multi-round auction mechanism where bidders submit their bids simultaneously on all licences each round. This multi-round process enables bidders to adjust their bids while taking into account the latest information about the likelihood of winning different sets of licences. Hence, at each round of the auction, a bidder can adapt its initial bidding strategy to the current market. This leaves room to a wide range of bidding strategies. Unfortunately, selecting the most efficient one is a difficult task which has not been solved yet due to the high complexity of the game induced by SAA [Reeves, 2005].

The aim of this section is to present how the bidding problem in SAA has been addressed in the literature until now. We start by defining the different features which characterise each bidder in SAA. Then, we present the complexities induced by the bidding game in SAA. We then introduce different metrics used throughout this thesis to evaluate the performance of a bidding strategy. We then present state-of-the-art bidding algorithms in SAA which are mainly used for benchmarking in this thesis. Finally, we discuss other relevant research on bidding in SAA and explain how our thesis contributes to the existing literature.

## 2.1 Bidder’s characteristics

In SAA, a bidder is defined by three features: its value function, its budget and its eligibility. The two first are independent of the auction’s mechanism and generally correspond to private information. The last is defined by the specific activity rules introduced in SAA. As the most common implementation of SAA is full transparency [Cramton, 2006], eligibility is generally public information. In all this thesis, we design by  $n$  the number of bidders and by  $m$  the number of items in a SAA. From now on, we will generally use the word item or good to design a licence as what follows only depends on the SAA mechanism and not on the exact nature of the items sold.

## 2.1.1 Utility and value function

Let  $L = \{1, \dots, m\}$  be a set of indivisible goods. Let  $\mathcal{P}(L)$  be the power set of  $L$ . The value function  $v_i : \mathcal{P}(L) \rightarrow \mathbb{R}$  of bidder  $i$  is defined as a function which maps each subset  $X \subset L$  to its monetised value  $v_i(X)$  for bidder  $i$ . For each set of items  $X$ , the utility  $\sigma_i(X, P) : \mathcal{P}(L) \times \mathbb{R}_+^m \rightarrow \mathbb{R}$  of bidder  $i$  is defined as the value  $v_i(X)$  minus the price of each item in  $X$ . For instance, at the end of the auction, if bidder  $i$  wins the set of items  $X$  and the bid price vector is  $P$ , then bidder  $i$  obtains a utility of:

$$\sigma_i(X, P) = v_i(X) - \sum_{j \in X} P_j \quad (2.1)$$

Value functions are assumed to be normalised ( $v_i(\emptyset) = 0$ ), finite and verify the free disposal condition, i.e., for any two sets of items  $X$  and  $Y$  such that  $X \subset Y$ , then  $v(X) \leq v(Y)$  [Lehmann, 2006; Milgrom, 2000]. The free disposal condition means that obtaining extra items does not generate any cost. They can be disposed for free. This implies that the value  $v(X)$  of a set  $X$  is necessarily positive. In practice, acquiring spectrum licences generally comes with a certain number of requirements such as deployment, coverage or speed obligations. Hence, due to these requirements, acquiring extra licences generates costs. However, the selling price and the intrinsic value of spectrum licences generally far outweighs such costs. Hence, the free disposal condition seems to be a reasonable assumption in the case of spectrum auctions. We define hereafter some widely used classes of value functions.

### 2.1.1.1 Superadditive value functions

**Definition 2.1. Superadditive value functions:** A value function  $v : \mathcal{P}(L) \rightarrow \mathbb{R}_+$  is said to be superadditive if for any set of items  $X, Y \in \mathcal{P}(L)$  with  $X \cap Y = \emptyset$ ,  $v(X \cup Y) \geq v(X) + v(Y)$ .

This class of value functions is generally used to study the effect of complementarities on the outcome of an auction. A bidder with a superadditive value function means that any set of items  $X$  exhibit complementarities with any other disjoint set of items  $Y$ . In other words, the value obtained by acquiring both disjoint sets of items  $X \cup Y$  is equal to the sum of values of each set plus a complementarity surplus for having obtained both sets. An extreme case is when each item in a set is worthless without the others. These items are known as *perfect complements*.

### 2.1.1.2 Subadditive value functions

**Definition 2.2. Subadditive value functions:** A value function  $v : \mathcal{P}(L) \rightarrow \mathbb{R}_+$  is said to be subadditive if for any set of items  $X, Y \in \mathcal{P}(L)$  with  $X \cap Y = \emptyset$ ,  $v(X \cup Y) \leq v(X) + v(Y)$ .

Subadditivity is the opposite of superadditivity in the sense that obtaining both disjoint sets of items is worth less than the sum of values of both sets. Hence, all sets are considered complement free. There exist different types of subadditivity. One of them is known as substitutability. The main idea behind substitutability is that raising the price of some items should not affect the demand on the other items. We present two classes of value functions exhibiting substitutability which are often used in auctions (and specifically in SAA) to obtain theoretical guarantees.

The first is a class of value functions which verifies the gross substitutes property [Lehmann, 2006]. It is based on the concept of preferred set at price  $P$ , i.e. a set  $X$  such that  $\sigma(X, P) = \max_{Y \in \mathcal{P}(L)} \sigma(Y, P)$ . The demand set at prices  $P$  is defined as the set of all preferred sets  $D(P) = \{X | \sigma(X, P) = \max_{Y \in \mathcal{P}(L)} \sigma(Y, P)\}$ .

**Definition 2.3. Gross substitutes property [Lehmann, 2006]:** A valuation  $v : \mathcal{P}(L) \rightarrow \mathbb{R}_+$  is said to satisfy the gross substitutes property if for any item  $j$ , any price vector  $P$  and any price vector  $P' \geq P$  component-wise with  $P'_j = P_j$ , the following implication is always true:

$$j \in X \text{ s.t } X \in D(P) \implies \exists Y \in D(P') \text{ s.t } j \in Y \quad (2.2)$$

The second is a class of value functions which verifies the mutual substitutes property. It was defined by Milgrom in [Milgrom, 2000]. It generalises the gross substitutes property in the sense that raising the prices of items not belonging to a set  $X$  does not reduce the demand of items in  $X$ . Hence, if items are mutual substitutes, then they are also gross substitutes.

**Definition 2.4. Mutual substitutes property [Milgrom, 2000]:** A valuation  $v : \mathcal{P}(L) \rightarrow \mathbb{R}_+$  is said to satisfy the mutual substitutes property if for any set of items  $X \in \mathcal{P}(L)$ , for every pair of price vectors  $P' \geq P$  component-wise with  $\forall j \in X, P'_j = P_j$ , the following implication is true:

$$X \subset Y \in D(P) \implies \exists Y' \in D(P') \text{ s.t } X \subset Y' \quad (2.3)$$

An extreme case of substitutability is perfect substitutes or *single-unit demand* [Gul, 1999] where  $\forall X \in \mathcal{P}(L), v(X) = \max_{j \in X} v(\{j\})$ . In other words, items are perfect substitutes for a bidder if it is only interested in the item with the highest utility in a set and disregards the rest.

### 2.1.1.3 Additive value functions

**Definition 2.5. Additive value functions:** A value function  $v : \mathcal{P}(L) \rightarrow \mathbb{R}_+$  is said to be additive if for any set of items  $X, Y \in \mathcal{P}(L)$  with  $X \cap Y = \emptyset$ ,  $v(X \cup Y) = v(X) + v(Y)$ .

In other words, a bidder admits an additive value function if it values any set of items as the sum of values of each item. Such value functions exhibit no complementarity and no substitutability. Items are considered as independent from the perspective of the bidder.

### 2.1.2 Budgets

The budget  $b_i$  of a bidder  $i$  defines the maximal amount that bidder  $i$  can spend in an auction. The final payment of bidder  $i$  can not exceed its budget  $b_i$ . In spectrum auctions, the budget is usually fixed by an operator's corporate management. Without loss of generality, budget  $b_i$  is chosen independently from the value function  $v_i$ . We make the classical assumption throughout this thesis that a bidder never bids above its budget  $b_i$ . Hence, if the current bid price vector is  $P$  in a SAA with bid increment  $\varepsilon$ , a bidder  $i$  temporarily winning a set of items  $Y$  can bid on a set  $X$  containing different items if and only if

$$\sum_{j \in X} (P_j + \varepsilon) \leq b_i - \sum_{j \in Y} P_j \quad (2.4)$$

This assumption eliminates any bidding strategy which includes *budget bluffing* [Porter, 2006]. This consists in bidding above one's budget to fool opponents into believing its budget is greater than it really is. This tactic is employed in the specific case where one is convinced that it will likely be outbid on some of the items. However, it is unclear what happens if one ends up by having to pay more than its budget and cannot afford to. Hence, such tactics are not considered in this thesis.



### 2.1.3 Eligibility

In addition to value functions and budgets, bidders also present eligibility points in SAA. These are introduced through its specific activity rules. In practise, the most common activity rules are monotonicity rules [Bichler, 2017] where, as the bid prices rise, a bidder's eligibility cannot increase. In this thesis, we decide to focus on the following monotonicity rule: the number of items temporarily won plus the number of new bids by a bidder, which corresponds to its eligibility, can never rise. This activity rule is often used in the literature due to its simplicity [Goeree, 2014; Milgrom, 2004].

More formally, if a bidder  $i$  is temporarily winning a set of items  $Y$  and bids on a set of different items  $X$  ( $X \cap Y = \emptyset$ ) at a given round, its eligibility  $e_i$  is defined as the number of items contained in each set, i.e.  $e_i = |X| + |Y|$ . In the next round, if bidder  $i$  is temporarily winning a set of items  $Y'$ , it can only bid on a set of different items  $X'$  ( $X' \cap Y' = \emptyset$ ) such that

$$|X'| \leq e_i - |Y'| \quad (2.5)$$

Its eligibility is then set to  $e'_i = |X'| + |Y'| \leq e_i$  (monotonicity). In addition to the bids submitted on  $X'$ , bidder  $i$  can also submit bids on items it is temporarily winning, i.e. it can continue bidding on any item in  $Y'$ .

In practise, the initial eligibility of a bidder is determined by the deposit it makes before the auction starts [Milgrom, 2000]. In spectrum auctions, this deposit represents the quantity of spectrum (in our case, the number of licences) it wishes to be eligible. To keep things simple, we set the initial eligibility of each bidder to the maximal number of items sold in the auction, i.e.  $\forall i \in \{1, \dots, n\}, e_i = m$ .

## 2.2 Complexities induced by the bidding game in SAA

Unlike truthful auctions, where bidding its value is known to be a weakly dominant strategy, no such results have been provided in Simultaneous Ascending Auctions in the general case. Presently, auction theory and exact game resolutions methods are unable to compute the optimal bidding strategy as the bidding game induced by SAA is too complex [Reeves, 2005]. Bidding in SAA presents three different types of complexities. The first are inherent to the class of game it belongs to, i.e. general characteristics of the game that defines it such as n-player games. We will refer to those as *general game properties*. The second known as *game complexity* defines the complexity of the game through different metrics.

We focus mainly on two metrics which are: *information set space complexity* and *game tree complexity*. Finally, *strategical complexities* which are caused by specific rules and the mechanism of SAA.

## 2.2.1 General game properties

Here are some of the general features describing the bidding game induced by SAA.

- **n-player game:** n-player games are games which are well defined for any number of players. In a SAA, any number of bidders can participate. The main difficulty compared to single player games, such as the card game Solitaire, is that defining optimality is not as straightforward. In single player games, one's expected utility only depends on the strategy played. No external constraints need to be taken in account in the maximisation process of one's expected utility. However, in n-player games, the expected utility returned by one's strategy also depends on the strategies played by its opponents. For instance, in Rock Paper Scissors, playing rock maximises my utility if my opponent is playing scissors but minimises it if my opponent plays paper. To overcome this issue, we will often rely on the concept of best response and Nash equilibrium [Myerson, 1978].
- **Non zero-sum game:** Unlike Poker or Chess, our bidding game is not a zero-sum game, i.e. a game in which for any strategy profile the sum of the players' utilities is equal to zero. Zero-sum games imply that if a player increases its utility by changing its strategy, then the utility of another player will decrease. As we assume throughout all this thesis that a bidder's utility only depends on the items won and their respective bid price, SAA induces a non-zero sum game. For example, in a SAA with only one item sold, if a bidder acquires the item at a bid price of 1 instead of 10, the sum of players' utilities increases by 9.
- **Simultaneous move:** During each round of a SAA, all bidders submit their bids simultaneously. This leads potentially to ties between bidders having bid the same amount on the same item. This would not have occurred in a sequential bidding game where bidders take turns bidding.
- **Stochasticity:** This is introduced by the random tie breaking rule which selects randomly a temporary winner amongst the players having bid the highest same

amount on a item. This implies that a same configuration of bids does not necessarily lead to the same next state.

- **Imperfect information:** Perfect information games are games where all players are aware of all their opponent's past moves. According to [Cramton, 2002], it is commonly accepted to reveal a maximum of information to favour competition. Hence, in this thesis, all bids and their holders are revealed at the end of each round. However, as bidding in SAA is a simultaneous move game, bidders bid without knowing the move played by their opponents during the same round. Hence, this induces a game with imperfect information. Moreover, one can imagine an SAA where each new bid is announced but not the identity of the bidder having submitted it. For example, the first trial of SAA did not announce bidder identities [Milgrom, 2000]. In this specific case, for any past bid leading to a temporary win of a licence, a bidder does not know which of its opponents has submitted it. Nevertheless, it is admitted that bidders are usually able to infer bidder identities anyway [Milgrom, 2000].
- **Incomplete information:** In a complete information game, all the game specificities such as rules, strategies, payoffs or budgets are common knowledge. Hence, as bidders do not know the exact value function and budget of their opponents in SAA, the induced bidding game is with incomplete information. We design by the term "type" a possible combination of one's private information. For instance, in a game with complete information, each player has only one type. However, if from the point of view of its opponents, a bidder  $i$  has either value function  $v_i^1$  or  $v_i^2$  and either budget  $b_i^1$  or  $b_i^2$ , then bidder  $i$  presents four types which are:  $(v_i^1, b_i^1)$ ,  $(v_i^1, b_i^2)$ ,  $(v_i^2, b_i^1)$  and  $(v_i^2, b_i^2)$ . In this thesis, we will start by supposing that the game is with complete information. This is motivated by the fact that, in spectrum auctions, large bidders have usually a relatively precise estimation of the value function of other large bidders as well as their budgets. We will then add later on uncertainty to the bidder's estimation of its opponents' value function and budget.

Most research has focused on 2-player zero-sum turn-based deterministic games with perfect and complete information such as Go, Chess or Shogi [Silver, 2017; Silver, 2016]. These games admit special and restrictive properties which makes them simpler to analyse mathematically [Maschler, 2020]. For instance, the form of the optimal strategy is given by Von Neumann's minimax theorem [Neumann, 1947]. What made these games so challenging to study is their gigantic size. For instance, the number of possible different states which can be legally reached in the game Go is  $10^{172}$  [Allis, 1994]. We describe, in the next chapter, some classical algorithms for solving such games. These algorithms

are a source of inspiration for the algorithms used in this thesis. It is important to note that a direct application of algorithms specific to 2-player zero-sum deterministic games with perfect and complete information is doomed to fail in our considered bidding game.

## 2.2.2 Game complexities

To measure the complexity of the bidding game, we focus essentially on two different metrics:

- **Information set space complexity:** This refers to the number of possible information sets which can legally be reached from the initial position of the game. An information set is a set of states which are indistinguishable for the concerned player at the current position of the game [Cowling, 2012a]. In imperfect information games, a player, given its current information, does not necessarily know at which exact state it is playing. It only knows that it is in one of the states belonging to the corresponding information set. This uncertainty does not occur in perfect information games where information sets are all singletons and, hence, only correspond to one state [Świechowski, 2023]. Thus, in perfect information games, information set space complexity is equal to state space complexity. Otherwise, information set space complexity acts as a lower bound of state space complexity. In this thesis, we denote the information set space by  $\mathcal{I}$  and the state space by  $\mathcal{S}$ .

It is common to represent simultaneous move games as sequential games where the actions of all players during a given round are only revealed at the end of this round [Browne, 2012]. Therefore, a state in a SAA game is defined by seven features: the identity of the next bidder to bid (we name it the "concerned bidder"), the eligibility of each bidder, the bid price of each item, the temporary winner of each item, the type of the concerned bidder, the types of the concerned bidder's opponents and bids already submitted during a given round but not yet revealed. The first five are common knowledge and the last two are hidden information for the concerned bidder at the given state. Thus, all states which only differ by the last two features belong to the same information set.

- **Game tree complexity:** This originally refers to the number of leaf nodes in the solution search tree of the game's initial state in a 2-player zero sum game [Allis, 1994]. The solution search tree of a node  $J$  is the full-width search tree of minimal depth needed to compute the game-theoretic value of node  $J$ , i.e. the utility vector

obtained at node  $J$  when both players play optimally. In many games such as Chess, computing exactly the game tree complexity is hardly feasible as, in addition to the high state complexity of the game, the number of legal moves and the depth needed to compute the game-theoretic value of a node  $J$  can vary drastically between different states. A common approximation is to raise the average number of moves per state to the power of the average game length. For instance, in chess, the average number of moves per state is 35 and the average game length is 80. Hence, its game tree complexity is considered to be  $10^{123}$  [Allis, 1994]. In one sense, the game tree complexity can be understood as the number of different paths in the full-width search tree starting from the initial state of the game with a maximal depth equal to the average game length. Hence, in this thesis, we will refer to the game tree complexity of our bidding game as the number of different possible scenarios ending after a given number of rounds.

The information set space complexity and game tree complexity of the game induced by SAA is huge. To give an idea of the order of magnitude of both complexities, let's take the SAA which took place in Italy in 2018 where 12 5G spectrum licences were sold between 5 telecommunication companies after 171 rounds [European 5G Observatory, 2018]. In the most simplified version of SAA game considered in this thesis (also named d-SAA, see Chapter 4), the information set space complexity and the game tree complexity for such values are respectively greater than  $10^{35}$  and  $10^{491}$ .

Today, no generic algorithm is known to solve n-player non zero-sum simultaneous move games with incomplete information with high information set space and game tree complexities [Russell, 2021].

### 2.2.3 Strategical complexities

To the difficulties generated by the general game properties and game complexities of SAA, a number of strategical issues adds up due to its specific rules and mechanism. The main strategical challenges are the exposure problem, own price effect, budget constraints and eligibility management. We detail these four issues below:

- **Exposure problem:** As already explained in Section 1.2.3, the exposure problem refers to the possibility that, by bidding on a set of complementary items, a bidder ends up paying more than its valuation for the subset it actually wins as the competition was tougher than expected. Hence, the exposed bidder obtains a negative utility. This

is highlighted by the following example, referred to as *Example 1*, which will be used repeatedly throughout the thesis.

**Example 1:** We consider a SAA with two items between two bidders with no binding budgets. Its bid increment  $\varepsilon$  is equal to 1. The first bidder considers both items as perfect substitutes, i.e. obtaining both items gives the same utility as just obtaining one of them. The second bidder considers both items as perfect complements, i.e. each item is considered worthless without the other. The values of each combination of items for each bidder are presented in Table 2.1.

	$v(\{1\})$	$v(\{2\})$	$v(\{1, 2\})$
Bidder 1	12	12	12
Bidder 2	0	0	20

**Tab. 2.1.:** Example of exposure ( $\varepsilon = 1$ ), also referred to as *Example 1*

*Example 1* is a famous example of exposure [Wellman, 2008]. Even though it seems relatively simple, it is considered as a challenging problem as most existing algorithms fail to return the optimal bidding strategy for the second bidder. To compute it, we need to anticipate how the first bidder is going to bid. The objective of the first bidder is to obtain one of the two items for a price lower than 12. Hence, a rational strategy for bidder 1 is:

- If temporarily winning no items and the cheapest item has a price weaker than  $12 - \varepsilon$ , to bid  $\varepsilon$  on the cheapest item.
- Otherwise, to not bid.

Given the fact that the first bidder plays rationally, the second bidder should immediately drop out of the auction. Indeed, if the second bidder decides to bid on both items during the first round, then it will end up temporarily winning one item at the end of the round. The first bidder will continue bidding until the second bidder decides to give away an item or both items have a price greater than  $12 - \varepsilon$ . In both cases, the second bidder ends up with a negative utility.

- **Own price effect:** Each new bid on a item increases its bid price and, hence, decreases the utility of all bidders wishing to acquire it. As bidders have a mutual interest of keeping bid prices low, they can form collusions [Brusco, 2002] and divide the items among themselves. Explicit communication between bidders is illegal.

Nevertheless, bidders can still anticipate their opponents' bidding strategy from their valuation estimates. From this prior, they can coordinate themselves on how to divide the items. Moreover, a bidder can decide to reduce unilaterally the number of items on which it bids in spite that its opponents will not raise the bid price of the items it is temporarily winning. This strategy where one concedes items to its opponents to avoid a rise in price is called demand reduction [Ausubel, 2014; Weber, 1997].

**Example 2:** We consider an SAA between two bidders with two items. Its bid increment  $\varepsilon$  is equal to 1. Both bidders have additive value functions, i.e. the value given to a package of items is equal to the sum of values of each individual item in the package. They value each item at 10.5. The bidders' value functions are presented in Table 2.2.

	$v(\{1\})$	$v(\{2\})$	$v(\{1, 2\})$
Bidder 1	10.5	10.5	21
Bidder 2	10.5	10.5	21

**Tab. 2.2.:** Example of own price effect ( $\varepsilon = 1$ )

Example 2 is a simple example where tackling own price effect has a huge impact on the auction's final outcome. We present below the two different outcomes in case of collusions or without colluding:

- **Collusion:** If both players decide to collude, then they both bid on one item during the first round. If they bid on two different items, then they both acquire an item for a price of  $\varepsilon = 1$ . If they bid on the same item, then the temporary loser bids on the other item during the second round. In both cases, both bidders obtain an item for a price  $\varepsilon = 1$  and a utility of 9.5.
- **Without collusion:** Both bidders try to obtain both items without ever conceding one item to their opponent. Hence, the bid price of each item rises. Bidders stop bidding when both items have a bid price of 10. Hence, both bidders obtain an expected utility of 0.5.

By colluding, each bidder increases its expected utility by 9. In other words, the expected utility obtained by colluding of each bidder is 19 times greater than if they had not colluded. This example highlights the fact that tackling the own price effect is essential in order to maximise one's utility.

- **Budget constraints:** In practice, no bidder has unlimited funds. Hence, the maximum amount that a bidder can spend during an auction is capped by a fixed budget. These caps can highly impact the auction's final outcome. For instance, they can prevent bidders from bidding on certain packages of items and be a source of exposure.

*Example.* A simple example where a bidder can end up exposed due to budget constraints is presented in Table 2.3. It is a SAA with 2 items between 2 bidders and a bid increment  $\varepsilon$  equal to 1. Similarly to *Example 1*, the first bidder considers both items as perfect substitutes while the second bidder considers both items as perfect complements. In the case of unlimited budgets, the second bidder should win both items and end up with a strictly positive utility. However, suppose now that bidder 1's budget  $b_1$  and bidder 2's budget  $b_2$  are both equal to 10. As already explained in *Example 1*, a rational strategy for bidder 1 is:

- If temporarily winning no items and the cheapest item has a price weaker than  $7-\varepsilon$ , to bid  $\varepsilon$  on the cheapest item.
- Otherwise, to not bid.

Given the fact that bidder 1 plays rationally, if bidder 2 decides to acquire both items, then it will end up exposed. Indeed, after the bid price of each item has reached a value of 5, bidder 2 can no longer bid on both items as its budget is too low. Hence, it ends up with only one item for a price of 5 and, therefore, obtains a utility of  $-5$ .

	$v(\{1\})$	$v(\{2\})$	$v(\{1, 2\})$
Bidder 1	7	7	7
Bidder 2	0	0	20

**Tab. 2.3.:** Example of exposure due to budget constraints ( $\varepsilon = 1, b_1 = 10, b_2 = 10$ ).

Foreseeing such issues, a bidder can drastically change its bidding strategy depending on its budget as well as on its estimation of its opponents' budgets.

Moreover, in order to obtain less competition on the items one wants, a bidder can bind its opponents' budgets by driving up the bid price of other items they are interested in. Hence, as its opponents are temporarily winning items with high bid price, they might no longer be able to afford bidding on the items that actually interests this specific bidder. This is known as *budget binding* [Bichler, 2017].



- **Eligibility management:** Efficient management of its own eligibility is a key factor to ensure a favourable outcome. By decreasing its eligibility too quickly, one might end up trapped in a vulnerable position where it can no longer bid on certain packages and, thus, no longer follow the bidding pace of the auction. However, if all bidders maintain a certain level of eligibility, then the bid price of each item increases each round. Hence, tackling the own price effect and maintaining a high eligibility are in opposition. In order to maximise one's utility, a tradeoff must be found between both strategical issues.

## 2.3 Performance indicators

In games, the most common and natural metric to measure the performance of a strategy is its *expected utility*. In 2 player zero-sum games, this is often referred to as the *winning rate* [Silver, 2016]. Other metrics which derive from it are also sometimes used such as the *Elo rating* [Coulom, 2008], initially used for chess ranking. Up until now, the *expected utility* has been the only performance indicator used to compare strategies in SAA [Wellman, 2008; Goeree, 2014]. This intrinsically supposes that bidders are risk-neutral, i.e. their only objective is to maximise their expected utility, and that they participate in a high number of auctions. However, a specific instance of a spectrum auction (i.e. same frequency bands, same operators, etc ...) is generally only held once and an operator just participates to a few different instances. Hence, comparing strategies only on the basis of their *expected utility* is not sufficient. Indeed, for example, taking into account the potential losses due to exposure is also relevant, especially considering the huge amount of money involved. In this thesis, we propose other indicators to measure the performance of bidding algorithms. In addition to the expected utility, we also focus on four other indicators (*expected exposure*, *exposure frequency*, *average price paid per item won*, *ratio of items won*) which quantifies the impact of *exposure* and the *own price effect*.

- **Expected utility:** The expected utility of a policy  $\pi$  depends on the strategies played by the other players. This is why game-theoretically the expected utility of a player is generally defined as a function of a profile of strategies. If the strategies of all opponents are fixed, the policy  $\pi^*$  which maximises the player's expected utility is named *best-response*. A profile of strategies where each player's strategy is a best response to the other players' strategies is known as a *Nash equilibrium* [Osborne, 2004]. These fixed points are of particular interest in auction theory (and in game theory in general) as none of the players can improve its expected utility by unilaterally

changing its strategy. Hence, the concepts of best-response and Nash-equilibrium are widely used in this thesis. However, it is important to note that a strategy belonging to a Nash equilibrium is only optimal if all the other players decide to play their respective strategy in the Nash equilibrium. Otherwise, playing such a strategy can perform poorly. Therefore, to partially avoid this issue, we also employ an empirical game analysis approach developed by Wellman et al. in [Wellman, 2008]. We compare two bidding algorithms by considering the normal form SAA game in expected utility where each bidder has the choice between either playing the strategy suggested by the first algorithm or the strategy suggested by the second. We then look at whether the strategy proposed by one of the two algorithms always strictly dominates the strategy proposed by the other, i.e. if a player has always interest in playing the strategy proposed by the same algorithm regardless of what the other plays.

- **Expected exposure:** It is based on the following decomposition of the expected utility:

$$\mathbb{E}(R^\pi) = \mathbb{P}(R^\pi \geq 0) \mathbb{E}(R^\pi | R^\pi \geq 0) + \underbrace{\mathbb{P}(R^\pi < 0) \mathbb{E}(R^\pi | R^\pi < 0)}_{\text{Exposure}} \quad (2.6)$$

where  $\pi$  is a policy and  $R^\pi$  is a random variable corresponding to the utility (generally named reward in reinforcement learning [Roijers, 2013]) obtained by playing  $\pi$  in a SAA. The second term in Equation 2.6 corresponds to the losses incurred when playing  $\pi$ . Hence, we define the expected exposure as  $-\mathbb{P}(R^\pi < 0) \mathbb{E}(R^\pi | R^\pi < 0)$ . In our extensive numerical experiments, the expected exposure is estimated by the opposite of all losses incurred by a strategy divided by the number of plays. For instance, if a bidder has played  $\pi$  in  $N$  auctions and has obtained the following utilities  $R_1^\pi, \dots, R_N^\pi$ , then the expected exposure of playing  $\pi$  is estimated by:

$$-\sum_{k=1}^N \frac{\mathbb{1}_{R_k^\pi < 0} R_k^\pi}{N} \quad (2.7)$$

- **Exposure frequency:** It is defined as  $\mathbb{P}(R^\pi < 0)$ . It corresponds to the probability of ending up exposed when playing a certain strategy  $\pi$ . In our extensive numerical experiments, we estimate it by the number of times a strategy incurs a loss divided by the number of times it is played. For instance, if a bidder has played strategy  $\pi$

in  $N$  auctions and has obtained the following utilities  $R_1^\pi, \dots, R_N^\pi$ , then the exposure frequency of playing  $\pi$  is estimated by:

$$\sum_{k=1}^N \frac{\mathbb{1}_{R_k^\pi < 0}}{N} \quad (2.8)$$

In order to tackle efficiently the exposure problem, both risk indicators, i.e. expected exposure and exposure frequency, should be minimised.

- **Average price paid per item won:** This is a relevant indicator to measure the impact of the own price effect as obtaining a low average price paid per item won means that the bid price of the purchased items rarely rises. In our extensive numerical experiments, we estimate it by the sum of expenses divided by the number of purchases induced by a strategy.
- **Ratio of items won:** To ensure that a strategy divides efficiently items between bidders and that no item is returned to the auctioneer unnecessarily, we consider the ratio of items won. Thus, the ratio of items won complements the preceding indicator for the measurement of the own price effect. In our extensive numerical experiments, we estimate it by the number of items won by a strategy divided by the number of items for sale in all auctions where it is played.

## 2.4 Benchmarking: perceived-price bidding strategies

Throughout this thesis, we are going to compare our bidding algorithms to state-of-the-art bidding strategies in SAA. More precisely, we are going to compare our bidding algorithm to a specific family of bidding strategies named *perceived-price bidding strategies* developed by Wellman et al. [Wellman, 2008]. This family of bidding strategies obtained very promising results but experiments were only undertaken with bidders having super-additive value functions, unlimited budgets and in SAA with no activity rules. Hence, it is difficult to conclude on how these bidding strategies perform when extended to more generic settings. We can define two types of perceived-price bidding strategies. The first type only depends on the current information set of the game. For example, this is the case of the well studied *straightforward bidding strategy* [Milgrom, 2000]. The second type also uses a prediction of closing prices, i.e., an estimation of the price of each item at the end of the auction. This prediction can be computed through different methods detailed hereafter.

In this subsection, we consider an SAA with  $n$  bidders,  $m$  items, a bid increment of  $\varepsilon$  and no activity rules. Bidders have unlimited budgets.

**Definition 2.1.** [Wellman, 2008] A perceived-price bidder is parameterised by a function  $\rho : \mathcal{I} \rightarrow \mathbb{R}_+^m$ , named perceived-price function, which maps an information set  $I \in \mathcal{I}$  to a perceived-price vector. When temporarily winning a set of items  $Y$  at current information set  $I$  with bid price vector  $P$ , it computes the subset of goods

$$X^* = \arg \max_{X \subset \{1, \dots, m\} \setminus Y} \sigma(X \cup Y, \rho(I)) \quad (2.9)$$

breaking ties in favour of smaller subsets and lower-numbered goods. The perceived-price bidder then bids  $P_j + \varepsilon$  on all items  $j \in X^*$ .

For example, a perceived-price bidder  $i$  is temporarily winning no item at information set  $I$  in a SAA with 2 items. If  $\sigma_i(\{1\}, \rho_i(I)) = \sigma_i(\{2\}, \rho_i(I)) = \sigma_i(\{1, 2\}, \rho_i(I)) > 0$ , then bidder  $i$  bids  $P_1 + \varepsilon$  on item 1 as it is the smaller subset and the lower-numbered good which maximises its perceived utility.

This family of bidding strategies is motivated by the following statement: if a bidder knew the closing prices of an auction and these prices are independent of its bidding strategy, i.e. these closing prices stay the same whatever the bidder decides to play, then playing as above using the actual closing prices as perceived-price vector is optimal [Wellman, 2008]. However, in practise, closing prices are generally tightly correlated to one's bidding strategy, especially regarding incumbents in spectrum auctions. Nevertheless, it is important to bear in mind that if the actual closing prices of a SAA can accurately be predicted, then a bidder will face no exposure problem, no problems allocating its limited budget and no eligibility management problem [Bulow, 2009].

### 2.4.1 Straightforward bidding

*Straightforward bidding* (SB) is surely the most studied bidding strategy in SAA. This name was given by Milgrom in [Milgrom, 2000]. The same solution concept is also referred to through other names in the literature such as "myopic best response" [Reeves, 2005].

**Definition 2.2.** A straightforward bidder temporarily winning a set of items  $Y$  at current information set  $I$  with bid price vector  $P$  is a perceived-price bidder with the following perceived-price function:

$$\rho_j(I) = \begin{cases} P_j & \text{if } j \in Y \\ P_j + \varepsilon & \text{otherwise} \end{cases} \quad (2.10)$$

In other words, a straightforward bidder maximises directly its utility given the items it is temporarily winning as if the auction ended immediately after its bid. It never takes into account the fact that other bidders are playing.

Even though this bidding strategy is naive, it has been proven to be optimal in a number of different situations. For instance, in a SAA with  $m = 1$ , SB is a weakly dominant strategy [Bikhchandani, 1997]. This can be easily understood by seeing that an SAA with  $m = 1$  is strategically equivalent to a sealed-bid second-price auction. Hence, SB in SAA with  $m = 1$  is similar to bidding truthfully in a sealed-bid second-price auction [Wellman, 2008]. Moreover, in a SAA where all bidders have a single-unit demand and value each item equally, i.e.  $v(\{1\}) = \dots = v(\{m\})$ , SB is a Bayes-Nash equilibrium [Wellman, 2008; Peters, 2006]. In the case of mutual substitutes, Milgrom shows in [Milgrom, 2000] that, if all bidders play SB, the final allocation approximately maximises the total value and every bidder acquires a set of items which is close to optimal with respect to the closing prices. Hence, the final allocation and closing prices represent an approximate competitive equilibrium.

In the general case, playing SB is far from optimal. For instance, in *Example 1* (see Section 2.2.3), playing SB for bidder 2 leads to exposure. Moreover, even in the case of additive value functions, SB is no longer a weakly dominant strategy due to the own price effect. For instance, in *Example 2* (see Section 2.2.3), consider the threat strategy where a bidder follows the collusion instruction the two first rounds and plays SB if its opponents deviates from these instructions. If both bidders play the threat strategy, they both obtain a utility of 9.5. If one of them decides to play SB, then they end up with a maximal utility of 1 depending on the outcome of the tie-breaking rule in the first round.

Nevertheless, in practise, playing SB can sometimes be a reasonable strategy. For instance, in the 3G UK SAA held in 2000 [Binmore, 2002], five licences, two bigger ones and three others, were sold with no bidders being allowed to buy more than one. Given these specific rules, according to [Bulow, 2009], playing SB is a good strategy and each licence will be allocated to the bidder willing to pay the most for it.

## 2.4.2 Point-price prediction strategy

**Definition 2.3.** [Wellman, 2008] A point-price prediction bidder temporarily winning a set of items  $Y$  at current information set  $I$  with bid price vector  $P$  is a perceived-price bidder whose perceived-price function uses an initial prediction of closing prices  $P^{init}$ :

$$\rho_j(I) = \begin{cases} \max(P_j^{init}, P_j) & \text{if } j \in Y \\ \max(P_j^{init}, P_j + \varepsilon) & \text{otherwise} \end{cases} \quad (2.11)$$

We denote this strategy throughout the thesis by PP.

As stated before, playing a perceived-price bidding strategy with the perceived-price vector initially equal to the actual closing prices of the auction is optimal. Hence, playing PP with  $P^{init}$  equal to the actual closing prices is optimal. Playing PP with  $P^{init} = 0$  is equivalent to playing SB. Hence, one can easily see that the efficiency of playing PP highly depends on the accuracy of one's initial prediction of closing prices  $P^{init}$ . For instance, if  $P^{init}$  largely underestimates the actual closing prices of the auction, then playing PP becomes equivalent to playing SB when  $P \geq P^{init}$  component wise. However, if  $P^{init}$  largely overestimates the closing prices of the auction, then a bidder playing PP might drop out prematurely of the auction as it erroneously anticipates exposure.

An easy way to compute an initial prediction of closing prices is to use the average final prices obtained when simulating a high number of times a fixed profile of strategies. In the case of incomplete information where the opponents' value function are unknown, one can sample their valuations from the underlying type distribution at each simulation [Wellman, 2008]. For instance, by sampling from the type distributions, one could use as initial prediction of closing prices the average final prices obtained when all bidders play SB. This would generally lead to an overestimation of the actual closing prices of the auction. Hence, playing PP with this initial prediction of closing prices would generally cause bidders to drop out early. A very popular prediction of closing prices often used to enable optimal allocation is the Walrasian equilibrium [Gul, 1999].

### 2.4.2.1 Walrasian equilibrium

**Definition 2.4.** [Blumrosen, 2007] A Walrasian equilibrium in an auction with  $n$  bidders is a price vector  $P$  and an allocation  $X_1, \dots, X_n$  such that

$$\forall i \in \{1, \dots, n\}, \sigma_i(X_i, P) = \max_{Y \in \mathcal{P}(L)} \sigma_i(Y, P) \quad (2.12)$$

For any unallocated item  $j$ , i.e.  $j \notin \cup_{i=1}^n X_i$ ,  $P_j = 0$  (market-clearing).

In the case of budget constraints, a Walrasian equilibrium  $(P, X)$  needs also to verify that  $\forall i \in \{1, \dots, n\}, \sum_{j \in X_i} P_j \leq b_i$ .

**Theorem 2.1.** [Blumrosen, 2007] *If the vector of prices  $P$  and the allocation  $X_1, \dots, X_n$  is a Walrasian equilibrium, then the allocation  $X_1, \dots, X_n$  maximises social welfare. This is a direct application of The First Welfare theorem to an auction with indivisible goods.*

It is important to note that the concept of Walrasian equilibrium does not take into account the auction mechanism. Therefore, given the above theorem, a Walrasian equilibrium  $(P, X)$  gives the optimal allocation in all auctions, i.e. it maximises  $\sum_{i=1}^n v_i(X_i)$ , and in particular in SAA. However, depending on the bidders' valuations, this equilibrium does not necessarily exist.

**Theorem 2.2.** [Kelso Jr, 1982] *An SAA where all bidders have their value function which verify the gross substitutes property admits a Walrasian equilibrium.*

**Theorem 2.3.** [Gul, 1999] *If a value function  $v_1$  does not satisfy the gross substitutes property, then there exist value functions  $v_2, \dots, v_n$  which satisfy the gross substitutes property such that no Walrasian equilibrium exist in SAA.*

It is highly plausible that no Walrasian equilibrium exist in SAA if not all value functions satisfy the gross substitute property as stated in Theorem 2.3. For example, in *Example 1* (see Section 2.2.3),  $v_1$  satisfies the gross substitutes property but  $v_2$  does not as it exhibits complementarities. In this example, there exist no price vector such that each bidder maximises its utility and the market clears. Hence, the SAA in *Example 1* does not admit a Walrasian equilibrium. Examples of auctions with two goods between two bidders where

one exhibit complementarities and the other substitutabilities are often used to show the absence of a Walrasian equilibrium [Wellman, 2008]. Hence, this solution concept can not be used generally in SAA.

Nevertheless, based on the idea of Walrasian equilibrium, Wellman et al. [Wellman, 2008] propose two methods for the initial prediction of closing prices: *expected price equilibrium* and *expected demand price equilibrium*.

#### 2.4.2.2 Expected price equilibrium

The aim of the expected price equilibrium is to compute the expectation (over opponents' type distribution) of the Walrasian equilibrium price vector. The algorithm works as follows:

1. Sample a value function for each opponent from their type distribution.
2. From these sampled value functions compute the demand function  $x : \mathbb{R}_+^m \rightarrow \mathbb{N}^m$  which maps for each item  $j$  a price vector  $P$  to the number of bidders willing to acquire it.
3. Let  $p(t)$  be the price vector at step  $t$  and  $\alpha(t)$  be an adjustment parameter which decays with  $t$ . For example, in [Cheng, 2003], an exponential decay is chosen for  $\alpha(t)$ . Initially,  $p(0)$  is equal to the null vector of prices. Then, apply the standard tâtonnement process [Arrow, 1971] to SAA with one unit of each good available:

$$p(t + 1) = p(t) + \alpha(t)(x(p(t)) - 1) \quad (2.13)$$

4. Repeat the sampling of the value functions and the application of the standard tâtonnement process to obtain an average vector of final prices.

At each iteration, value functions are sampled from the opponents' type distributions and then the tâtonnement process in Equation 2.13 is applied. This tâtonnement procedure is guaranteed to converge to a Walrasian equilibrium if the sampled value functions satisfy the gross substitutes property [Cheng, 2003].

It is important to note that this method does not depend on the auction mechanism. It only depends on the type distributions and adjustment parameter  $\alpha(t)$ . Hence, in a SAA with



complete information, if all bidders compute the expected price equilibrium using the same adjustment parameter  $\alpha(t)$ , then they obtain the same initial prediction of final prices.

We denote by EPE the bidding strategy PP using as initial prediction of final prices the expected price equilibrium.

### 2.4.2.3 Expected-demand price equilibrium

Expected-demand price equilibrium is an alternative to expected-price equilibrium which might be preferred for computational reasons. Instead of sampling from the type distribution and applying the tâtonnement process many times, one could compute the expected demand function and then apply once and for all the tâtonnement process. In other words, the expected-demand price equilibrium is the simple application of the standard tâtonnement process in Equation 2.13 using the expected aggregate demand. The demand function no longer returns in  $\mathbb{N}^m$  but in  $\mathbb{R}_+^m$ .

We denote by EDPE the bidding strategy PP using as initial prediction of final prices the expected-demand price equilibrium. It is important to note that, in a complete information game, the expected-demand price equilibrium returns the same price vector as the expected price equilibrium if computed with the same adjustment parameter  $\alpha(t)$ . Hence, in a complete information game, playing EPE or EDPE is equivalent.

### 2.4.3 Distribution price prediction strategy

Instead of just having a point estimate of the final prices of an auction, one could decide to use an entire distribution. Let  $\mathcal{D}$  be a distribution predictor of the final prices of the auction. Let  $\mathbf{P}(p|I)$  be the probability according to  $\mathcal{D}$  that the final prices of the auction will be  $p$  given that the current information set is  $I$ . Let  $I_0$  be the initial information set of the game. Before the auction starts, given our initial prediction, the probability that the closing prices of the auction are equal to  $p$  is  $\mathbf{P}(p|I_0)$ . Then, at information set  $I$  with current bid price vector  $P$ , the probability that the closing prices of the auction are equal to  $p$  is:

$$\mathbf{P}(p|I) = \begin{cases} \frac{\mathbf{P}(p|I_0)}{\sum_{q \geq P} \mathbf{P}(q|I_0)} & \text{if } p \geq P \\ 0 & \text{otherwise} \end{cases} \quad (2.14)$$

where  $p \geq P$  means that  $\forall j \in \{1, \dots, m\}, p_j \geq P_j$ . This is well defined for all possible  $P$  if only  $\exists p \geq P, \mathbf{P}(p|I_0) > 0$ . Therefore, it is common [Wellman, 2008] to put a small probability weight on an upper-bound of  $P$  to ensure that Equation 2.14 is always well defined. Let  $\bar{P}$  be an upper-bound of  $P$ . This can easily be computed by using the type distributions or the budget distributions.

Final prices between different items are generally correlated. However, for sake of simplicity, a common approach is to consider them as independent and, thus, use the vector of marginal distributions  $(\mathcal{D}_1, \dots, \mathcal{D}_m)$ . Moreover, Wellman et al. [Wellman, 2008] assume that the bidder is fully sunk aware, i.e. that the perceived-price of temporarily won items that it expects to win at the current bid price is equal to zero. They introduce the concept of expected incremental price for an item  $j$  at information set  $I$  that we denote by  $\Delta_j(I)$ . It corresponds to the expected sunk cost of winning item  $j$ . To compute  $\Delta_j$ , two disjoint cases must be studied, depending on whether a bidder is temporarily winning item  $j$  or not.

If a bidder is not winning an item  $j$  at information set  $I$  with current bid price  $P$ , then the lowest price at which it can buy item  $j$  is  $P_j + \varepsilon$ . Hence, its expected incremental price for item  $j$  is:

$$\Delta_j^L(I) = \mathbb{E}_{\mathcal{D}}(p_j | p_j \geq P_j + \varepsilon) = \sum_{q_j \geq P_j + \varepsilon} \mathbf{P}(q_j | p_j \geq P_j + \varepsilon) q_j \quad (2.15)$$

If a bidder is temporarily winning an item  $j$  at information set  $I$  with current bid price  $P$ , we need to consider two cases to compute its expected incremental price. If the final price of item  $j$  ends up by being  $P_j$ , then the full price is sunk as the item is already committed. Otherwise, there is  $1 - \mathbf{P}(P_j | p_j = P_j)$  chance that the bid price rises. In this case and with the additional assumption that a bidder never bids on the items it is temporarily winning, the current winner will have to bid at least  $P_j + 2\varepsilon$  to win back item  $j$ . Hence, its expected incremental price for item  $j$  is:

$$\Delta_j^W(I) = (1 - \mathbf{P}(P_j | p_j = P_j)) \sum_{q_j \geq P_j + 2\varepsilon} \mathbf{P}(q_j | p_j \geq P_j + 2\varepsilon) q_j \quad (2.16)$$

A distribution price prediction bidder then plays a perceived-price bidding strategy at information set  $I$  using as perceived-price function:

$$\rho_j(I) = \begin{cases} \Delta_j^W(I) & \text{if winning item } j \\ \Delta_j^L(I) & \text{otherwise} \end{cases} \quad (2.17)$$

### 2.4.3.1 Self-confirming price distribution

The concept of self-confirming price distribution was originally introduced for SAA [Wellman, 2008]. It has then been used in other auctions such as simultaneous one-shot sealed bids auctions [Wellman, 2012].

**Definition 2.5.** [Wellman, 2008] Let  $\Gamma$  be an instance of a SAA game. A distribution of prices  $\mathcal{D}$  is said to be self-confirming if  $\mathcal{D}$  is the distribution of final prices resulting when all bidders play the distribution price prediction strategy using  $\mathcal{D}$  as price distribution predictor in  $\Gamma$ .

As it was the case for the distribution price prediction strategy, even though there are generally dependencies across prices, we consider the marginal price distributions as if goods were independent.

**Definition 2.6.** [Wellman, 2008] Let  $\Gamma$  be an instance of a SAA game. The vector of marginal price distribution  $(\mathcal{D}_1, \dots, \mathcal{D}_m)$  is said to be self-confirmed if  $\forall j \in \{1, \dots, m\}$ ,  $\mathcal{D}_j$  is the marginal distribution of final prices of item  $j$  when all bidders play the distribution price prediction strategy using  $\mathcal{D}$  as price distribution predictor in  $\Gamma$ .

To compute a self-confirming vector of marginal price distributions, Wellman et al. [Wellman, 2008] propose to use a fixed-point iteration method:

1. Let  $\mathcal{D}^t$  be the vector of marginal price distributions at step  $t$  (for  $\mathcal{D}^0$  the uniform distribution is generally chosen between 0 and the bid price upper bound  $\bar{P}$ ).
2. Run many game instances and record their resulting final prices. For each instance, sample the value distributions from their type distribution and then simulate a SAA game where all bidders play the distribution price prediction strategy using  $\mathcal{D}^t$  as price distribution predictor.
3. For each item  $j$ ,  $\mathcal{D}_j^{t+1}$  is set such that the probability that  $p$  is the final price of  $j$  is equal to the number of times  $p$  ended up as the final price of  $j$  divided by the number of simulations.

No guarantees are provided for the convergence of this process. Self-confirming price distribution do not always exist as it is the case in *Example 1* (see Section 2.2.3). When an

approximate fixed point is not found, a solution which is proposed by [Osepayshvili, 2012] is to average the distributions obtained over the last few steps.

We denote by SCPD the distribution price prediction bidding strategy which uses the above fixed-point iteration method to derive its marginal price distribution predictor. In case of non-convergence of the method, we average the distributions obtained in the last ten iterations as in [Osepayshvili, 2012].

## 2.5 Other related works

The majority of studies on SAA, such as [Cramton, 2006; Cramton, 2002; Milgrom, 2000], focuses on its mechanism design, efficiency and the revenue it generates for the regulator. Only a few works in the literature have focused on the strategic bidding problem in SAA. At first glance, this can seem very surprising, especially considering the fact that SAA is the most popular mechanism used for spectrum auctions and sometimes billions of euros are at stake. This may be due to the fact that such expertise is very valuable and, hence, operators or auction consulting firms have no interest in sharing their own private research on the subject.

Until now, published works on the bidding problem have mainly focused on trying to resolve one of the four strategical issues (exposure, own price effect, budget constraints and eligibility management problem) in specific simplified versions of SAA. These strategical issues are generally studied separately. For example, to the best of our knowledge, no efficient bidding algorithm had been proposed before our contributions [Pacaud, 2022] to tackle simultaneously the own price effect and the exposure problem. Moreover, most solutions can often only be applied to small instances and, therefore, can generally not be used in practise. In this subsection, we propose an overview of the different methods used in the literature to tackle each strategical issue.

### 2.5.1 Exposure

Exposure is considered as the most problematic strategical issue in SAA. It has essentially been studied in two formats of SAA: its original format [Cramton, 2006] described in Section 1.3.2.1 (this is the format considered in this thesis) and its corresponding clock format defined hereafter.

In the original format of SAA, from the regulator's point of view, the fact that the exposure problem reduces the auction's efficiency and revenue has been highlighted by Milgrom [Milgrom, 2000] and by Bykowsky et al. [Bykowsky, 2000] in small examples of SAA with complete information. From the bidder's point of view, an interesting approach, which has already been presented in the last section, is perceived-price bidding strategies [Wellman, 2008]. By using a prediction of final prices, bidding strategies such as EPE, EDPE or SCPD can improve one's expected utility by tackling efficiently the exposure problem. Experimental results suggest that these methods are very promising. However, they were only obtained in the case of unlimited budgets, no activity rules and for a specific class of super-additive value functions. No theoretical guarantees were provided with these methods. Hence, it is difficult to conclude on the effectiveness of these algorithms in more generic settings.

The original format of SAA is generally considered too complex to draw theoretical guarantees. Hence, a simplified clock format of SAA [Goeree, 2014] is often considered. Under some specific conditions, it presents the advantage of being a tractable model where bidders have continuous and differentiable expected utilities. Standard optimisation methods can then be applied to derive an equilibrium.

The clock format of SAA presents two types of bidders: *local* and *global*. Local bidders only compete in one of the English separate auctions while global bidders participate in all auctions. Local bidders play straightforwardly, i.e. bid up to their value, as it is a weakly dominant strategy. Global bidders have super-additive value functions and, therefore, are prone to exposure. In this specific format, the price of each item is modelled by a clock that ticks upward at equal and constant pace if, at least, two bidders accept the current price level. If only one bidder accepts the current bid price of an item, the price clock pauses and the bidder becomes its temporary winner. It employs the same activity rule as the one used in this thesis, i.e. the number of items temporarily won plus the number of new bids by a bidder can never rise.

In this format, Goeree and Lien [Goeree, 2014] use a Bayesian framework to compute the global bidder's optimal drop-out level in homogeneous-good environments. The fact that all items are identical is essential for tractability as it enables a one-dimensional type approach, i.e. the value function of a bidder only depends on the number of items won and no longer on specific packages. They extend their work to two specific cases. The first case is a clock-format SAA with two global bidders with regional complementarities, i.e. global bidders are only interested in a fixed number of identical items. The second case authorises local bidders to switch from one English auction to another. In this case, local bidders consider all items as perfect substitutes. This will cause bid prices to even out.

Zheng proposes in [Zheng, 2012] to modify the initial clock format of SAA by integrating a pause system that enables jump bidding. Jump bidding is the practise of bidding considerably more than the minimal admissible bid. They then build a continuation equilibrium that fully eliminates the exposure problem in the specific case of an SAA with two items and one global bidder. More precisely, let's suppose there are two auctions: auction 1 and auction 2. If all local bidders drop out of auction 1, the global bidder is given the opportunity to jump bid in auction 2 to see if it can obtain the second item at a profitable price. If some local bidders decide still to stay in auction 2 after the jump bid, the global bidder can then decide to quit both auctions to avoid exposure. It then concedes item 1 to the local bidder who triggered the pause. The global bidder can also decide to just to drop out of auction 2 if winning item 1 is profitable. Otherwise, the global bidder wins both items. It pays item 2 the same amount as its jump bid. By modifying the initial SAA mechanism as above, it is possible to build a continuation equilibrium that completely eliminates the exposure problem in very small instances.

## 2.5.2 Own price effect

Own price effect has mainly been studied in the original format of SAA. In [Milgrom, 2000], Milgrom describes a *collusive equilibria* in a 2-object SAA with complete information between two bidders having additive value functions. The analysis is similar to the one given in *Example 2* (see Section 2.2.3).

Milgrom's work was then pursued by Brusco and Lopomo [Brusco, 2002] who build a collusive equilibrium in a 2-object SAA with incomplete information between two bidders in the case of additive value functions through signalling of the most valuable item. More precisely, during the first round, each bidder bids on its most valuable item. If bidders bid on different items, then they both obtain their most valuable item at the minimal price. Otherwise, each round, the temporary loser of the item bids the minimal admissible amount on the item which maximises its utility until one of them bids on the other item. Obviously, this collusive equilibria only holds if, given their opponent's value distribution, bidders are better off colluding than competing for the items in expectation. They then show that the probability of building such equilibria where bidders collude through signalling decreases when the number of bidder increases. In other words, the scope of collusion narrows when the ratio between the number of bidders and the number of object increases. They then extend their work to the case of a 2-object SAA between two bidders with superadditive value functions. They show that high complementarities is not the main factor that impacts collusion. For instance, in the complete information case where both bidders have the same

value function, bidders should collude. However, if the value for both items is sufficiently different between both bidders, then collusion may no longer be attractive.

Riedel and Wolfstetter [Riedel, 2006] show, given a certain number of constraining assumptions, that the only Nash equilibrium that survives iterated elimination of dominated strategies is the profile of strategies where bidders immediately reduce their demand to the efficient allocation. Thus, the auction ends after the first round. More precisely, they consider a clock-format of SAA with complete information, no budgets, no activity rules and identical goods. At each round, bidders state simultaneously the number of items they would like. The auction ends when the sum of demanded items is inferior to the total number of items. Bidders have strictly decreasing marginal valuations and all marginal valuations must be distinct between bidders. For instance, this implies that the marginal valuation of obtaining a third item for bidder 1 must be distinct from the marginal valuation of obtaining a second item for bidder 2. In this specific case, if the efficient allocation assigns at least one item to each player and the bid increment is sufficiently small, then the unique equilibrium that survives iterated elimination of dominated strategies is the one where bidders reduce their demand in the first round according to the efficient allocation.

Weber analyses in [Weber, 1997] how mutual reduction between bidders can be arranged to tackle the own price effect in the FCC spectral auction of 1994 even if bidders can not communicate between each other directly. For instance, an implicit agreement can be found by threatening to raise the prices of the licences held by the violator. Such threats to deter aggressive bidding and encourage collusions are often used in the literature, especially in game theory, to form collusive equilibria.

Using a similar approach than for the exposure problem with perceived-price bidding strategies and predictions of final prices, Wellman et al. [Wellman, 2008] build a simple bidding algorithm in a homogeneous-good environment where all players have subadditive value function. However, experimental results were unsatisfactory as they are significantly inferior to ones obtained with a simple demand reduction strategy.

### 2.5.3 Budget

In [Brusco, 2009], Brusco and Lopomo focus on "noncollusive" equilibria in a 2-object clock-format of SAA between two global bidders with super-additive value functions. Noncollusive equilibria are defined as equilibria where bidders decide to split items only when their

budgets are binding. Hence, the collusion is directly caused by the budget constraints. Value functions are private but budgets are common knowledge.

Brusco and Lopomo consider a clock-format of SAA with a pause-system which presents some important differences with the one proposed by Zheng [Zheng, 2012]. Each bidder pushes initially two buttons corresponding to the two items. When a button is released, it can not be pushed again. Unlike [Zheng, 2012], they have a single price and buttons are not object-specific, i.e. both items will be sold at the end of the auction at the same price. Their pause-system is triggered the first time one of the two bidders releases a button. More precisely, suppose bidder 1 releases two buttons. Then, the price stops raising and bidder 2 has either the choice between releasing its two buttons or winning both items at the current price. If it releases both buttons, then the entire package of two items are assigned to one of the bidders randomly. If bidder 1 releases just one button, then bidder 2 has the choice between releasing two buttons, one button or zero buttons. If it releases both buttons, bidder 1 wins both items. If it releases one button, each bidder wins one item. If it decides to release zero buttons, the price clock resumes. In this case, the price will raise until bidder 1 releases its last button which means bidder 2 wins both items or bidder 2 releases a button which means each bidder wins an item.

Using this specific clock-format, Brusco and Lopomo show the effect of budgets constraints on the structure of noncollusive equilibria. Moreover, they present various inefficiencies due to the combined effect of the exposure problem and budget constraints. One of these inefficiencies may appear surprising. Suppose bidder 1 has a lower budget and a lower value for the package than bidder 2. They exhibit some situations where bidder 1 ends up by winning both items as bidder 2 is afraid that bidder 1 is going to continue to bid on a single object. Hence, to avoid potential exposure, bidder 2 drops out of the auction.

In [Bulow, 2009], Bulow et al. propose a similar approach to avoid exposure than the perceived-price bidding strategies defined by Wellman et al. [Wellman, 2008] in the case of binding budget constraints. More precisely, their main motive is that the exposure problem and the issues linked to budget constraints could be eliminated if one could predict accurate final auction prices. However, unlike [Wellman, 2008] and traditional auction theory, their forecasting approach does not focus on bidders' values but on bidders' budgets. They define the *bid exposure* as the sum of all bids placed by a bidder during a given round, including its temporary winning bids from the prior round. By tracking a bidder's bid exposure, one will get a finer estimate of a bidder's real budget and eventually, after a few rounds, the bidder's bid exposure will approximately correspond to its real budget. More precisely, if all bidders have a binding budget constraint and start by bidding on many items, then one



could get a pretty precise estimate of the sum of all budgets quite early on in the auction. Then, at a certain point in the auction, the bid exposure of each bidder should stay relatively the same and bidder should start bidding on smaller subset of items. Combining the bid exposure of a bidder with the smaller subset of items on which it is bidding should give a relatively good idea of the final price of each item in the auction. Hence, if final prices are forecast sufficiently early, one can easily avoid exposure.

#### 2.5.4 Eligibility management problem

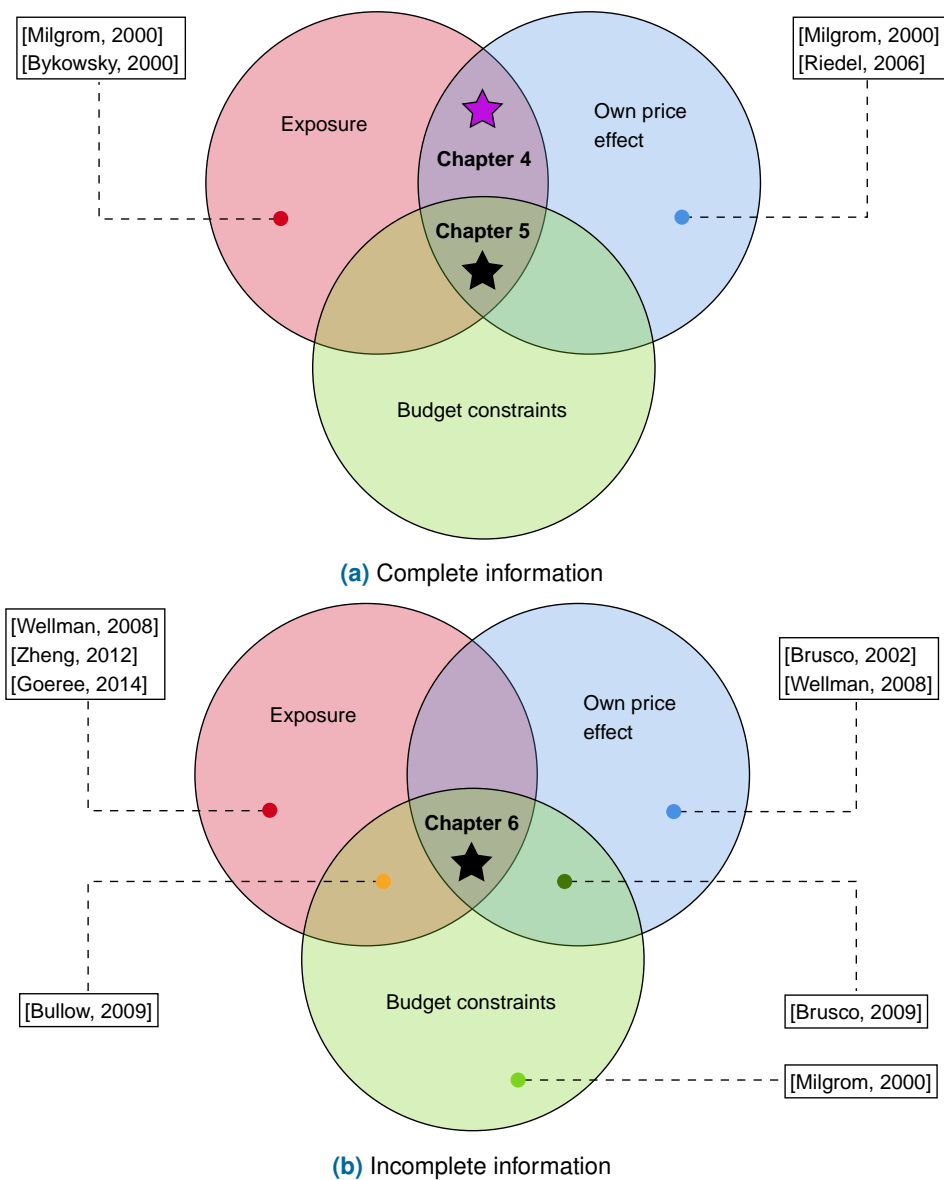
Even if some simplified versions of the SAA present activity rules, little work has been done regarding the eligibility management problem, especially in the original format of SAA. However, it is commonly accepted that one should gradually reduce its eligibility to avoid being trapped in a vulnerable position if other bidders do not behave as expected [Weber, 1997]. Hence, to maintain high eligibility and avoid rising the price of the items one is interested in, a bidder can decide to bid on highly demanded items it does not want. Moreover, this can also fool other bidders which now think one desires these highly demanded items. This practise is known as *parking* [Porter, 2006].

## 2.6 Positioning of the thesis

We present in Figure 2.1 how our thesis fits in the literature. More precisely, we focus on which strategical issues are covered in the different chapters of this thesis as well as in the related works. We only consider the three following strategical issues: exposure problem, own price effect and budget constraints in Figure 2.1 as the eligibility management problem has not really been addressed in the literature. Nevertheless, this last issue will be tackled in Chapter 5 and Chapter 6 of our thesis. Moreover, we make the distinction between bidding with complete information and bidding with incomplete information. Related works and chapters of our thesis represented in Figure 2.1(a) only address an SAA with complete information. However, all related works and Chapter 6 represented in Figure 2.1(b) can of course be applied to an SAA with complete information as it is a specific case of incomplete information. Moreover, we like to emphasise on the fact that typical examples taken from the literature are often with complete information as it is easier to determine if the proposed strategy tackles correctly the targeted strategical issue.

In Chapter 4, we propose the first algorithm that tackles efficiently and simultaneously the exposure problem and own price effect in small instances of a deterministic turn-based

version of SAA with complete information. In Chapter 5, we propose the first algorithm that tackles efficiently the four main strategical issues of SAA in its original format with complete information and on instances of realistic size. In Chapter 6, we relax the assumption that the induced bidding game is with complete information and generalise the algorithm presented in the preceding chapter to the incomplete information framework.



**Fig. 2.1.:** Strategical issues addressed by related works and by the different chapters of our thesis in SAA.



# A survey on adversarial search

## Contents

---

3.1	Extensive form . . . . .	64
3.2	Game tree and search tree . . . . .	67
3.3	Adversarial search in two player zero-sum games . . . . .	68
3.3.1	Minimax search . . . . .	68
3.3.2	$\alpha\beta$ search . . . . .	70
3.3.3	Expectimax search . . . . .	71
3.4	Adversarial search in n-player games . . . . .	72
3.4.1	Max <sup>n</sup> search . . . . .	72
3.4.2	Other classical adversarial search methods . . . . .	73
3.5	Adversarial search in games with large state space . . . . .	74
3.5.1	Heuristic evaluation functions . . . . .	75
3.5.2	Monte Carlo evaluation . . . . .	76
3.6	Bandits . . . . .	76
3.7	Monte Carlo Tree Search . . . . .	79
3.7.1	Selection . . . . .	80
3.7.2	Expansion . . . . .	81
3.7.3	Rollout . . . . .	82
3.7.4	Backpropagation . . . . .	83
3.8	Related works on MCTS . . . . .	83
3.8.1	Application of MCTS to two-player deterministic zero-sum games with perfect and complete information . . . . .	83
3.8.2	Application of MCTS to n-player games . . . . .	84
3.8.3	Application of MCTS to non-deterministic games with chance events . . . . .	84
3.8.4	Application of MCTS to simultaneous move games . . . . .	84
3.8.5	Application of MCTS to imperfect information games . . . . .	85
3.8.6	Application of MCTS to games with incomplete information . . . . .	88
3.8.7	Applications of MCTS to auction games . . . . .	88
3.9	Transposition table . . . . .	88
3.10	Conclusion . . . . .	90

---

In this thesis, our aim is to compute an efficient bidding strategy in SAA. Given the fact that SAA can be modelled as a multi-round game between many players with potentially conflicting goals, we need to be able to plan a bidding strategy over many rounds while taking into account how the other players are going to react. Hence, a natural approach is to focus on adversarial search methods.

Adversarial searches correspond to different search methods used in sequential games where two or more players have conflicting goals. Their aim is to compute the best possible move at a given state by exploring the state space.

In this chapter, we first start by defining theoretically a game in extensive form. We then define two important notions which arise from this representation: game tree and search tree. These are used as the basis of all adversarial search methods presented in this thesis. We then give an overview of different adversarial searches addressed in the literature depending on general game properties and game complexities. We start by presenting famous adversarial search methods used in two player zero-sum games which are by far the most studied class of game in the literature. We then present some possible extensions of these methods to n-player games. We address the problem of efficient adversarial searches in games with large state space and game tree complexity. We introduce the multi-armed bandit problem, considered as one of the most fundamental reinforcement learning problems, and present the famous *UCB1* and *EXP3* algorithms. We finally introduce Monte Carlo Tree Search (MCTS), a breadth-first search method, on which are based our solutions throughout this thesis. We then focus on different versions of MCTS presented in the literature, especially those applied to games sharing similar general game properties than our SAA game. Some of these algorithms present enhancements that we incorporate in our solution. Moreover, we specifically detail one common search enhancement named transposition tables.

## 3.1 Extensive form

In this subsection, we define formally a game in extensive form. This representation is based on the notion of tree.

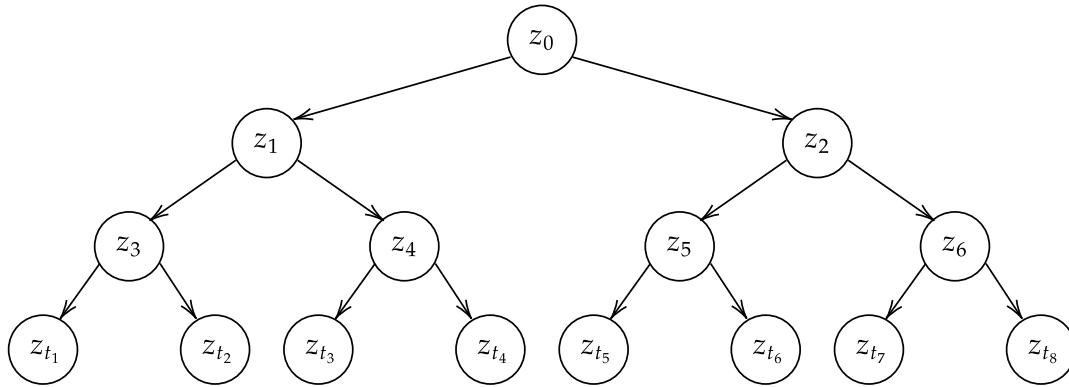
**Definition 3.1.** [Lasaulce, 2011] A tree is a triplet  $(Z, z_0, \theta)$  where:

- $Z$  is a set of nodes.
- $z_0$  is the root of  $Z$
- $\theta$  is the predecessor function. All nodes are linked to the root  $z_0$ , i.e.  $\forall z \in Z, \exists k \in \mathbb{N}, \theta^{(k)}(z) = z_0$ . There is no cycle.

Terminal nodes of a tree are nodes which do not have any successors. We design by  $Z_T$  the terminal nodes of a tree. More formally,

$$z_t \in Z_T \implies \forall z \in Z, \nexists k \in \mathbb{N}^*, \theta^{(k)}(z) = z_t \quad (3.1)$$

In Figure 3.1, we plot an example of a tree of depth 3.



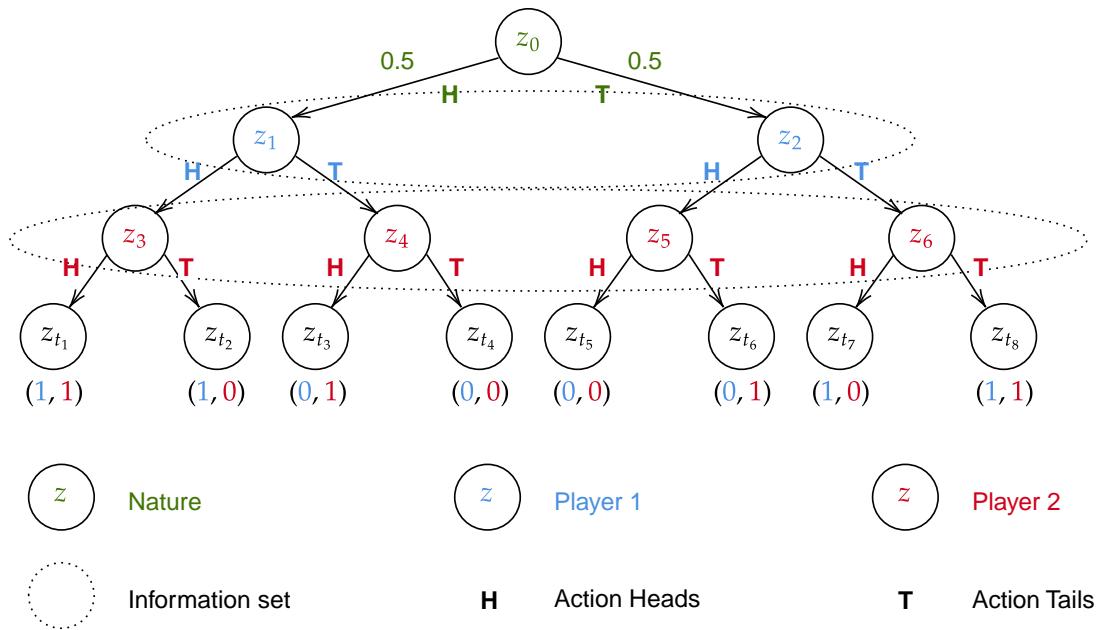
**Fig. 3.1.:** An example of a tree

**Definition 3.2.** A game in extensive form is defined by:

- A finite number of players  $n$
- A tree  $(Z, z_0, \theta)$
- A partition  $\{Z_i, i \in \{1, \dots, n\}\} \cup Z_{Nature}$  of  $Z \setminus Z_T$ . If  $z \in Z_i$ , then it is player  $i$  who plays at node  $z$ . If  $z \in Z_{Nature}$ , then it is a player named Nature who plays at node  $z$ . The player Nature is introduced to model chance events.
- $\forall z \in Z \setminus Z_T$ , a set of actions  $A(z)$ .

- $\forall z \in Z_{Nature}$ , a probability distribution  $\mathcal{D}(z)$  is defined over  $A(z)$ .
- Each player set  $Z_i$  is partitioned into subsets named information sets (already defined in the previous chapter). If  $z$  and  $z'$  belong to the same information set  $I$ , then  $A(z) = A(z')$ .
- A utility function  $\sigma : Z_T \rightarrow \mathbb{R}^n$

Each node of a game's extensive form represents a state of the game and each edge a possible action (either played by a player or drawn by Nature). We represent in Figure 3.2 a game in extensive form between two players using the same tree as in Figure 3.1. The game works as follows. Two players are standing on the same side of a wall. A coin is tossed and lands on the other side of the wall. None of the players know the outcome. There is 50% chance that the outcome is heads and 50% that the outcome is tails. Each player needs to guess the outcome. They write their answers on a sheet of paper without letting the other player know what they have written. Then, they both climb the wall to check the coin. If a player guesses correctly, then it wins and obtains a utility of 1. Otherwise, it obtains a utility of 0.



**Fig. 3.2.:** An example of a game in extensive form

This representation is often used as a starting point by adversarial search methods, such as minimax search presented hereafter, to compute an efficient strategy for the root player.

Although in the rest of the thesis we do not redefine each quantity systematically, this representation underlies our methods and is quite representative of the structure of our code.

## 3.2 Game tree and search tree

In order to compute the best possible move at a given state, adversarial search methods are based on the exploration of the game's state space. One possible approach is to consider the game's extensive form and compare the outcomes obtained in all different sequences of moves, i.e. all different paths of the tree. However, exploring the whole state space is practically impossible in games such as chess. Thus, we define the two following notions: game tree and search tree.

**Definition 3.1.** The game tree represents the actual game rooted at the current state of the game. More precisely, it is the tree associated to the subgame starting from the current state of the game. Thus, the sequence of moves leading to the current state of the game is not represented in the game tree. Each path from the root node corresponds to a possible sequence of actions of the game and ends at a terminal state. At each move played in the real game, we redefine the game tree as the subtree of the old game tree starting at the new current state.

*Remark.* The game tree rooted at the initial state of the game is the same tree as the one in the game's extensive form representation.

Exploring the whole game tree in games with huge state space (or information state space) is practically impossible. Hence, adversarial search methods generally focus only on a small portion of the game tree named search tree.

**Definition 3.2.** A search tree is the portion of the game tree on which the adversarial search is performed. Different statistics are saved for each node of the search tree. It shares the same root as the game tree.

A search tree is either constructed by a set of deterministic rules such as cutting off the game tree at a certain depth or it can be constructed dynamically by an adaptive process as



it is the case in Monte Carlo Tree Search presented in Subsection 3.7. Each leaf node of the search tree is given a value by either an evaluation function or by Monte Carlo Evaluation. The adversarial search method is then performed on this search tree to choose the best move.

## 3.3 Adversarial search in two player zero-sum games

Many algorithms have been created for adversarial search in two player games, especially in zero-sum games. Zero-sum games correspond to games where the sum of payoffs of the players at the end of the game is equal to zero [Lasaulce, 2011]. In a two player game, this means that a player's gain is equivalent to its opponent's loss. The basic approach for such games is to use the minimax algorithm presented in Subsection 3.3.1. However, this algorithm naively explores all the nodes of the game tree. It is possible to limit the exploration to fewer nodes using  $\alpha\beta$  pruning which we present in Subsection 3.3.2. The above algorithms are guaranteed to compute the optimal strategy in two player deterministic zero-sum games with perfect and complete information. According to Zermelo theorem [Zermelo, 1913], in this specific class of games, either one of the players has a winning strategy regardless of what its opponent plays or both players can guarantee themselves a draw. However, many games also present chance events or hidden events. For such games, a variant using chance nodes can be applied called expectimax described in Subsection 3.3.3.

### 3.3.1 Minimax search

Minimax search is the classical depth-first search method for sequential two-player zero-sum games. It is based on Von Neumann's minimax theorem [Neumann, 1947] and, hence, guaranteed to return the optimal move if applied to the whole game tree in a two player zero-sum deterministic game with perfect information. The two players are called *Max* and *Min*. *Max* player is the player playing at the current state of the game while *Min* player is its opponent. For games having a small state space such as Tic-Tac-Toe, the search tree used at each step is generally the whole game tree. However, if the state space is too big, it is practically impossible to explore the whole game tree. Thus, only a small portion of the game tree is used as the search tree. A common approach is to consider the subtree containing all nodes of the game tree until a certain depth  $d$ . The algorithm performs a depth-first exploration of the search tree to find the optimal move for *Max* player to play. The time complexity of such an algorithm is  $O(b^d)$  and space complexity is  $O(bd)$  with  $b$  the

average branching factor of the game, i.e. the average number of moves per state. A value will be assigned to the terminal nodes of the search tree corresponding to the *Max* player's payoff. For example, a win for the *Max* player could be rewarded by 1, a loss by  $-1$ , a draw by 0 and an undetermined issue (if the terminal node in the search tree is not a terminal node in the game tree) by 0.

The principle of this method is the following: *Max* player tries to maximise its payoff and the *Min* player tries to minimise it. We recall that, as we are studying a two player zero-sum game, the fact that *Min* tries to minimise *Max* payoff is equivalent to *Min* trying to maximise its own payoff. A node corresponding to *Max* player's turn will be assigned the maximum value amongst its children whereas a node corresponding to *Min* player's turn will be assigned the minimum value amongst its children. The minimax algorithm backtracks these values from the terminal nodes of the search tree to its root using recursion. The algorithm then returns the move leading to the root node's child with the highest value. An example is provided in Figure 3.3 for the game Tic-Tac-Toe.

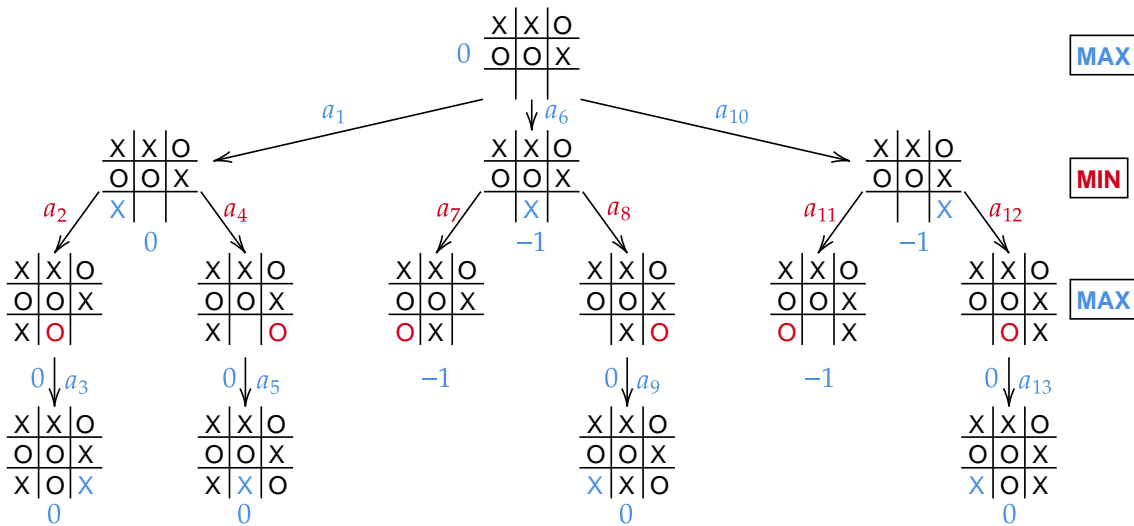


Fig. 3.3.: An example of a minimax search in Tic-Tac-Toe

In this example, the *Max* player is represented by the crosses and the *Min* player by the draughts in the Tic-Tac-Toe game. Values for move  $a_1$ ,  $a_6$  and  $a_{10}$  are backtracked recursively from the terminal nodes of the game. As action  $a_6$  and  $a_{10}$  leads to a loss while action  $a_1$  leads to a draw for *Max* player, the minimax algorithm returns action  $a_1$ . These actions are not named randomly but are named after their order of passage in this depth-first search method.

### 3.3.2 $\alpha\beta$ search

Exploring every node in the minimax search tree to compute the value of the root node and return the best move is unnecessary. Some branches of the search tree can be completely disregarded if they are determined as suboptimal early on. This method of eliminating branches to avoid exploring uselessly is called pruning.  $\alpha\beta$  pruning [Knuth, 1975] is one of the most commonly pruning techniques applied in minimax trees.

$\alpha\beta$  pruning is a technique that keeps in memory two bounds  $\alpha$  and  $\beta$  for each explored node for which the algorithm hasn't yet determined its value. At any node,  $\alpha$  corresponds to the minimal value and  $\beta$  to the maximal value that *Max* player can obtain from this node given the search performed so far. The real value of the node is in the interval  $[\alpha, \beta]$ . Thus, from these bounds, it is possible to identify if a branch is suboptimal without fully exploring it and, hence, eliminate it. The effectiveness of such a technique is highly dependent on the order in which the states are examined. Therefore, an ordering function is often implemented to guide the exploration of nodes. An optimal ordering function enables the  $\alpha\beta$  algorithm to examine the best successors first. In this best case, the time complexity is dropped to  $O(b^{\frac{d}{2}})$  whereas it was  $O(b^d)$  for minimax. For example in chess, the simple ordering function that tries to capture first, then tries to avoid threats, then tries to forward moves and finally tries backward moves gets close to  $O(b^{\frac{d}{2}})$  [Russell, 2021]. If a random move ordering to the  $\alpha\beta$  is applied then the time complexity of the algorithm is dropped to  $O(b^{\frac{3d}{4}})$ .

Figure 3.4 shows a simple example of how the  $\alpha\beta$  algorithm proceeds with no ordering function, i.e. nodes are explored in the same order as the minimax algorithm. At the beginning of the game, no guarantees are given on the value of root A. Thus, the real value of A is between  $\alpha_A = -\infty$  and  $\beta_A = +\infty$ . The algorithm then explores node B. As B isn't a terminal node, then  $\alpha_B = -\infty$  and  $\beta_B = +\infty$ . The next node is a terminal node which has for value 5. This value is backtracked recursively to node B. As B is a *Min* node, then 5 is assigned to  $\beta_B$  which means that the maximum value that *Max* player can get given the current information if it moves to the B node is 5. The two other children of B are explored. Therefore, as all nodes of subtree with root B have been examined, the real value of B is known and is equal to the minimum value of its children which is 5. Knowing that if *Max* player plays B, its payoff will be of 5, 5 is assigned to  $\alpha_A$ . The next node C is explored and its first child has value 0 so  $\beta_C = 0$ . *Max* player knows that, if it moves to C, its maximum payoff will be 0 whereas if it moves to B its payoff will be 5 ( $\beta_C \leq \alpha_A$ ). Therefore, there is no point in exploring subtree with root C anymore and pruning is applied on the rest of the branches of this subtree. The last node D highlights the fact that the move ordering function

is not optimal as  $\alpha\beta$  is obliged to explore all its children before concluding that moving to D is suboptimal for *Max* player.

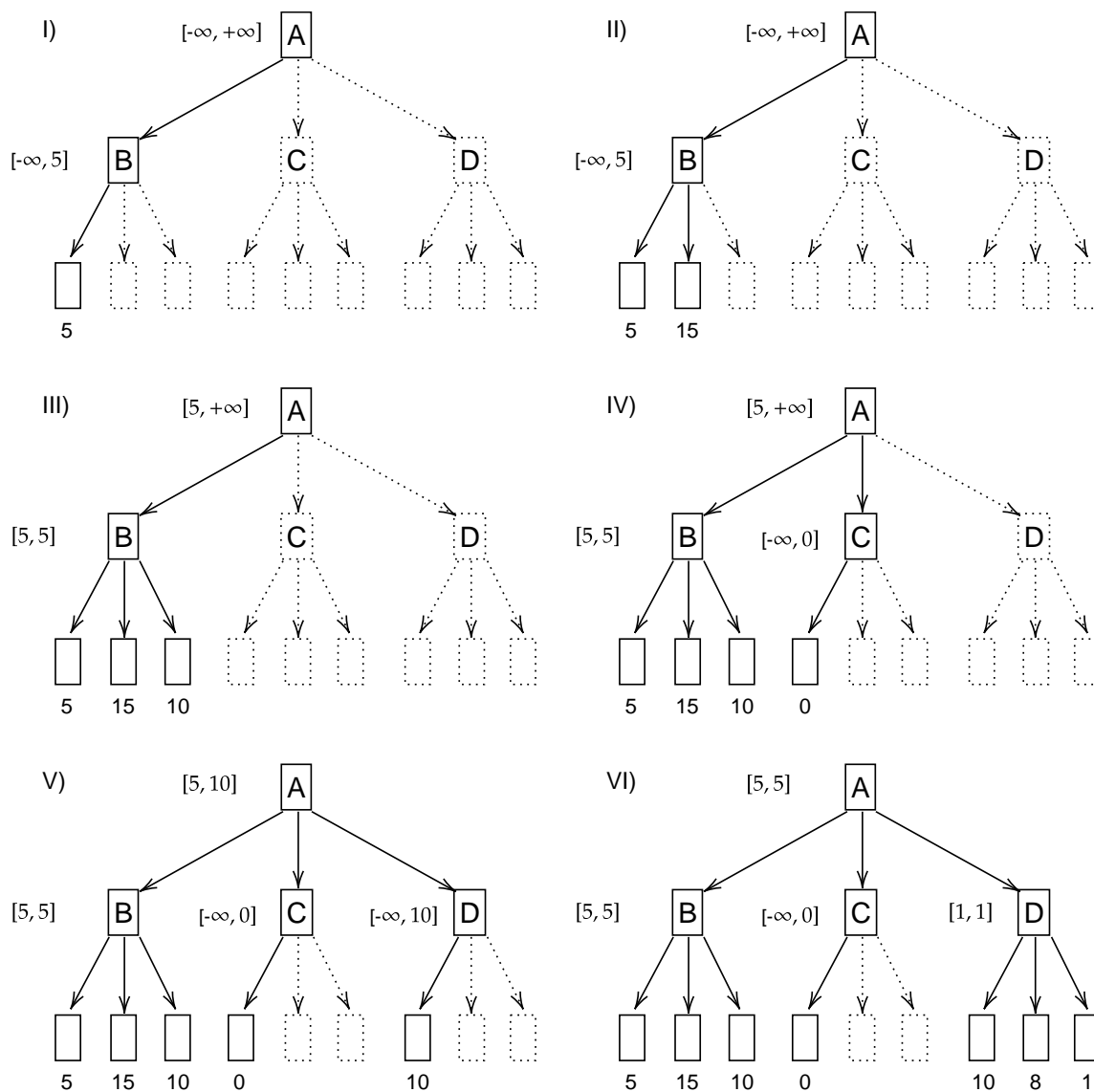


Fig. 3.4.: An example of a  $\alpha\beta$  pruning

### 3.3.3 Expectimax search

The minimax and  $\alpha\beta$  algorithms described as above can not deal with games where luck is an integral part. These are games such as backgammon where rolling a dice determines the actions of a player at each round or poker where a dealer deals cards. In the game's extensive form, these events are represented by the actions drawn by Nature according to

a certain probability distribution. If  $\forall z \in Z_{Nature}$ , the probability distribution  $D(z)$  is known, then these events can be represented in the search tree by *chance nodes*. The value of a chance node will then be defined as the weighted average of the value of its children.

$$V_{c^*} = \sum_{c \in C^*} \mathbf{P}(c) V_c \quad (3.2)$$

where  $c^*$  is a chance node,  $C^*$  is the set of children of  $c^*$ ,  $\mathbf{P}(c)$  is the probability that child  $c$  will be chosen according to the corresponding distribution  $\mathcal{D}(c^*)$  and  $V_c$  is the value of node  $c$ .

Thus, by adding chance nodes to the search tree, two player games presenting chance events with known probabilities can be dealt with using the same process as the minimax algorithm. This new algorithm is named expectimax [Michie, 1966]. However,  $\alpha$  and  $\beta$  can't be computed at chance nodes. There exist other pruning techniques to avoid exploring the whole expectimax search tree [Schadd, 2009].

## 3.4 Adversarial search in n-player games

The above algorithms can only be applied in two player zero-sum games. However, many games are played by more than two players. We refer to these games as *n-player games*. In this section, we briefly describe basic adversarial search methods applied in n-player games. Games considered hereafter are no longer necessarily zero-sum games.

### 3.4.1 $\text{Max}^n$ search

$\text{Max}^n$  [Luckhart, 1986] is probably the most popular technique used in n-player games. The outcome obtained at each leaf node of the search tree is a vector of size  $n$ , equal to the number of players. The values are backtracked recursively to the root node. At each node, the vector which is backpropagated is the vector that maximises the score of the player playing at this specific node. Thus, in a two player zero-sum game,  $\text{max}^n$  search is equivalent to minimax search.

Figure 3.5 illustrates the backpropagation process of  $\text{max}^n$  search in a three player non-zero sum game. Moreover, this example highlights an important problem concerning the backpropagation in case of ties. Indeed, player B has the choice between backpropagating  $(2, 6, 4)$  and  $(3, 6, 3)$ . Depending on the tie breaking rule,  $\text{max}^n$  may return a different

solution. A common tie breaking rule is to choose to backpropagate the vector which has the lowest value for the root player. If ties still remain, then the vector is selected randomly amongst the remaining vectors. This is the tie breaking rule used in Figure 3.5.  $\alpha\beta$  pruning can not be applied with  $\max^n$  but other pruning methods exist [Sturtevant, 2000]. This algorithm is easily extended to n-player games presenting moves by Nature with known probabilities by introducing chance nodes.

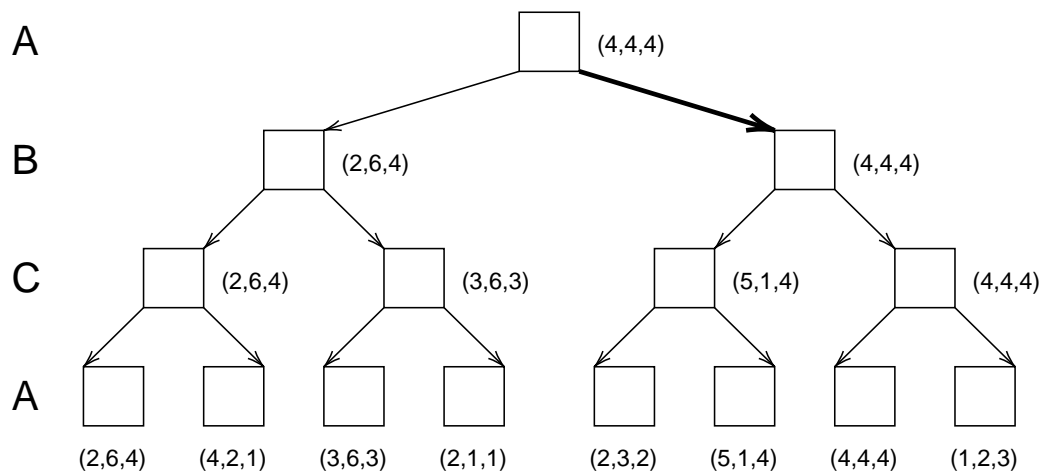


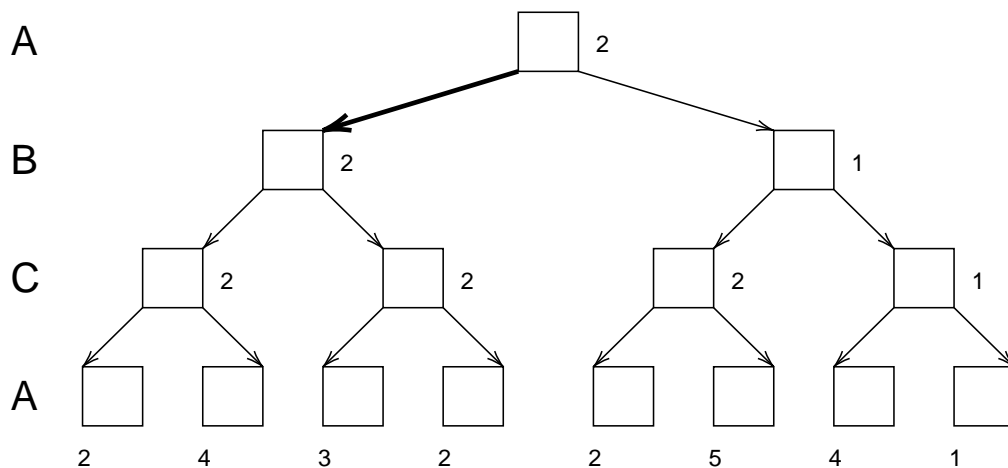
Fig. 3.5.: An example of  $\max^n$  search

### 3.4.2 Other classical adversarial search methods

Two other interesting adversarial search methods in n-player games are *Paranoid search* and *Best reply search* as they both use a minimax search tree. This enables them to use  $\alpha\beta$  pruning and, hence, spend less time exploring the search tree.

*Paranoid search* [Sturtevant, 2000] is based on the assumption that all opponents team up against the root player. Thus, under this assumption, the search tree can be represented as a minimax search tree where the *Max* player corresponds to the root player and the *Min* player to its opponents.

In Figure 3.6, we have plotted the paranoid search in the same game displayed in Figure 3.5. The game is reduced in a two player zero sum game using the utility of the root player as only outcome at the leaf nodes of the search tree. The value of the game, defined in Section 2.2.2, at the root node is then backtracked recursively using the minimax algorithm.



**Fig. 3.6.:** An example of paranoid search

One of the main disadvantages of paranoid search is that, as the assumption that all opponents form a coalition against the root player is usually false, the player tends to play too defensively. Indeed, as paranoid search tries to find the best solution in the worst-case scenario, more optimistic search algorithms usually compute better strategies.

To overcome this, *Best-Reply search* [Schadd, 2011] provides a less pessimistic view by allowing only one opponent to make a move instead of all opponents. This algorithm also reduces the  $n$ -player game in a two player zero-sum game using the utility of the root player as only outcome at the leaf nodes of the search tree. For each possible move of the root player in the search tree, the best counter amongst its opponents is selected. Thus, any sequence of moves in the search tree alternates between a decision taken by the root player and the best counter played by one of its opponents. The minimax algorithm is then applied to this minimax search tree. This adversarial search method gives the opportunity to explore more *Max* nodes and, hence, is able to compute longer term strategies than  $\max^n$  search and Paranoid search. However, the best counter selected can sometimes lead to illegal or unreachable states from the current state of the game. This can considerably reduce the relevance of the best move proposed by Best-reply search.

### 3.5 Adversarial search in games with large state space

One of the biggest issues when dealing with games with large state space is that the leaf nodes of the search tree are rarely terminal nodes of the game. This is problematic for the

backpropagation process of the aforementioned algorithms as the game-theoretic values for such states are generally unknown. To overcome such an issue, two standard methods are generally used: *heuristic evaluation functions* and *Monte Carlo evaluation*. These methods assign values to non-terminal nodes of the game in order to enable the backpropagation process and the use of adversarial search methods.

### 3.5.1 Heuristic evaluation functions

Heuristic evaluation functions are estimators of the expected final outcome of a game given its current state. Most evaluation functions consist in calculating various features of a state and return a linear combination of these features as estimation of the final outcome of the game. For example, a common heuristic evaluation function in chess is to calculate the number of different pieces of each colour, i.e. the number of white pawns, the number of black pawns, the number of white rooks, etc. The algorithm then returns a linear combination of these features. The different feature weights of the linear combination are generally either tuned by hand or by using different gradient methods. The choice of the features is purely empirical. The combination of  $\alpha\beta$  search and complex heuristic evaluation function achieved great success in the late 90s. For instance, the chess machine Deep Blue, developed by IBM, defeated the World Chess Champion Gary Kasparov in 1997. Its evaluation function is based on the recognition of 8000 patterns. A specific value is assigned to each pattern. These patterns and their corresponding value were nearly all created/tuned by hand, requiring a ton of expertise. Given a board's position, the evaluation function essentially returns the sum of values of the recognised patterns [Campbell, 2002].

To avoid such biased choices, neural networks taking the whole game state can be used as evaluation functions. These neural networks can be trained by using different reinforcement learning or supervised learning methods. One of the most famous adversarial search algorithm which uses a neural network as evaluation function is *AlphaZero* [Silver, 2017]. It is an improvement of author's preceding algorithm *AlphaGo* [Silver, 2016] which defeated one of the best Go players in the world, Lee Sedol, in 2016. By updating its weights through self-play, AlphaZero has mastered games such as Go, Chess or Shogi.

It is important to note that the performance of an adversarial search method strongly depends on the quality of its evaluation function. Indeed, as the values backpropagated to the root node of the search tree are principally computed by the evaluation function, the final selected best move by the adversarial search method highly depends on the evaluation function.



### 3.5.2 Monte Carlo evaluation

Constructing an efficient evaluation function can be complex. An alternative to estimate the expected utility of a non-terminal state is to return the averaged outcome of several random game plays from this state. This method is called Monte Carlo evaluation [Abramson, 1990]. More precisely, each simulation, also known as *rollout*, starts from the non-terminal state which is being evaluated. During the rollout, each player plays a random move. When the rollout reaches a terminal state, the final utility obtained by all players is computed and saved. After a certain number of rollouts, the Monte Carlo evaluation returns the average utility obtained by each player as estimation of the expected final outcome of the non-terminal state it is evaluating.

Monte Carlo evaluation achieved great success in deterministic games such as Go. Indeed, applying such a method with a minimax tree of depth 1 gave promising results in [Bouzy, 2004]. Moreover, Monte Carlo evaluation can also be applied to non-deterministic games such as backgammon [Tesauro, 1996].

The main drawback of Monte Carlo evaluation is that it is time-consuming. Indeed, to have sufficiently precise statistics at each leaf nodes of the search tree, many rollouts need to be performed. Thus, to apply Monte Carlo evaluation, the search tree needs to be of relatively small depth.

## 3.6 Bandits

Performing the same number of rollouts for each leaf node of the search tree is one of the major source of inefficiencies in adversarial search methods using Monte Carlo evaluation. Indeed, each evaluation is time-consuming and not all leaf nodes need to have a high number of rollouts before concluding that the action leading to it is suboptimal. A solution to this issue was provided by the *multi-armed bandit problem* [Lattimore, 2020]. This problem can be formulated as follows. Suppose there is a set of  $K$  slot machines and, thus,  $K$  arms to pull. An unknown distribution  $R_k$  of mean  $\mu_k$  is attributed to each slot machine  $k$ . The player decides to pull an arm  $k$  each turn and obtains a reward drawn from the unknown distribution  $R_k$ . Its aim is to maximise its cumulative reward and, thus, to figure out as quickly as possible which slot machine has the highest mean  $\mu_k$ .

In its simplest form, rewards are assumed to be independent and identically distributed. Each reward distribution is stationary. This is known as *stationary stochastic multi-armed bandit problem*.

As all multi-armed bandit problems, it introduces an important issue known as the *exploration-exploitation dilemma*. Exploitation consists in selecting the arm which seems to have the highest mean given the data observed so far. By doing so, the player tries to ensure itself a certain reward level. However, such a decision can prevent the player from finding better options. Exploration consists in selecting another arm by assuming that the data collected so far is not sufficient to identify the arm with the highest mean. Selecting such an arm and obtaining a bad reward will just reinforce the assumption that this arm is suboptimal. This will also result in a cost for the player. However, if this arm obtains a high reward, this could potentially mean that it is a better solution than the current arm used for exploitation. Thus, by potentially discovering a better solution through exploration, the player might considerably increase its future cumulative reward than if it had just stuck to pulling the arm proposed by the prior exploitation phase.

One of the most popular ways to solve the *stationary stochastic multi-armed bandit problem* is to use the *Upper Confidence Bound* algorithm (UCB) [Lattimore, 2020]. This method computes an upper confidence bound of each arm's mean. UCB is described in Algorithm 1.

---

#### Algorithm 1 UCB

---

**Input:** The exploration factor  $C$

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

- If an arm hasn't yet been selected then pull this arm.
- Otherwise pull arm  $k$  that maximises

$$\bar{x}_k + C \sqrt{\frac{\log(n)}{n_k}}$$

with  $\bar{x}_k$  the average reward obtained for arm  $k$  and  $n_k$  the number of times arm  $k$  has been pulled.

- Update  $\bar{x}_k$  with the obtained reward  $x_k$ .

**end for**

---

The term  $\bar{x}_k$  encourages the algorithm to select arms with high empirical mean while the term  $\sqrt{\frac{\log(n)}{n_k}}$  encourages the algorithm to select arms which have not been pulled often. The

tradeoff between exploitation and exploration is controlled by the coefficient  $C$ . One way of selecting the exploration factor  $C$  is to choose one which satisfies *Hoeffding's inequality*.

**Definition 3.1.** *Hoeffding's Inequality* [Lattimore, 2020]: If  $Z_1, \dots, Z_N$  are independent random variables almost surely in  $[a, b]$  then for all  $\delta \in [0, 1]$ , we have:

$$\mathbb{P} \left( \sum_{t=1}^N Z_t - E \left( \sum_{t=1}^N Z_t \right) \geq (b - a) \sqrt{\frac{N}{2} \log \left( \frac{1}{\delta} \right)} \right) \leq \delta \quad (3.3)$$

Under the assumption that the rewards are in  $[0, 1]$ , it can be shown that the *Hoeffding's inequality* gives us  $C = \sqrt{2}$  for an optimal tradeoff. This specific UCB algorithm is called *UCB1* [Auer, 2002a].

One famous extreme variant of the *multi-armed bandit problem* is the *adversarial multi-armed bandit problem* [Auer, 1995]. No statistical assumptions are made about the generation of rewards as it was the case for the *stationary stochastic multi-armed bandit problem*. At each turn, an adversary selects a reward function which maps each arm  $k$  to a specific value  $r_k$ . Simultaneously, the player pulls an arm  $k$  and receives reward  $r_k$ . The adversary is considered adaptive in the sense that it can learn a player's strategy. Given its knowledge of a player's strategy, it selects the worst-case reward function. Hence, having a deterministic strategy as it is the case with the *UCB1* algorithm can easily lead to poor performances. Thus, in order to maximise its cumulative reward, a player must maintain a probability distribution over the arms. Then, each turn, the player samples an arm according to this distribution and pulls it.

One of the most popular and simplest approaches to solve the *adversarial multi-armed bandit problem* is to use the EXP3 algorithm [Auer, 1995], which stands for "exponential-weight algorithm for exploration and exploitation". It is based on mixing two different distributions: a uniform distribution and a Gibbs distribution. The first distribution is used for exploration and the second for exploitation. EXP3 is fully described in in Algorithm 2:

By dividing reward  $x_k$  by the probability of selecting arm  $k$ ,  $s_k$  acts as an estimate of the sum of rewards for arm  $k$  over all iterations  $n$  and not just over the iterations where arm  $k$  is selected. The tradeoff between exploitation and exploration is controlled by two hyperparameters:  $\gamma$  and  $\eta$ . According to [Auer, 2002b], a good tradeoff between exploitation and exploration can be found for  $\gamma = \min(1, \sqrt{\frac{K \ln(K)}{(e-1)n}})$  and  $\eta = \frac{\gamma}{K}$  where  $e$  is the base of the natural logarithm.

---

**Algorithm 2** EXP3

---

**Inputs:** Hyperparameters  $\gamma$  and  $\eta$

**Initialisation:** Set  $s_k = 0$  for  $k = 1, \dots, K$

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

- Select arm  $k$  with probability  $p_k$ :

$$p_k = \frac{\gamma}{K} + (1 - \gamma) \frac{e^{\eta s_k}}{\sum_{l=1}^K e^{\eta s_l}}$$

- Receive reward  $x_k$
- Update  $s_k = s_k + \frac{x_k}{p_k}$  for selected arm  $k$

**end for**

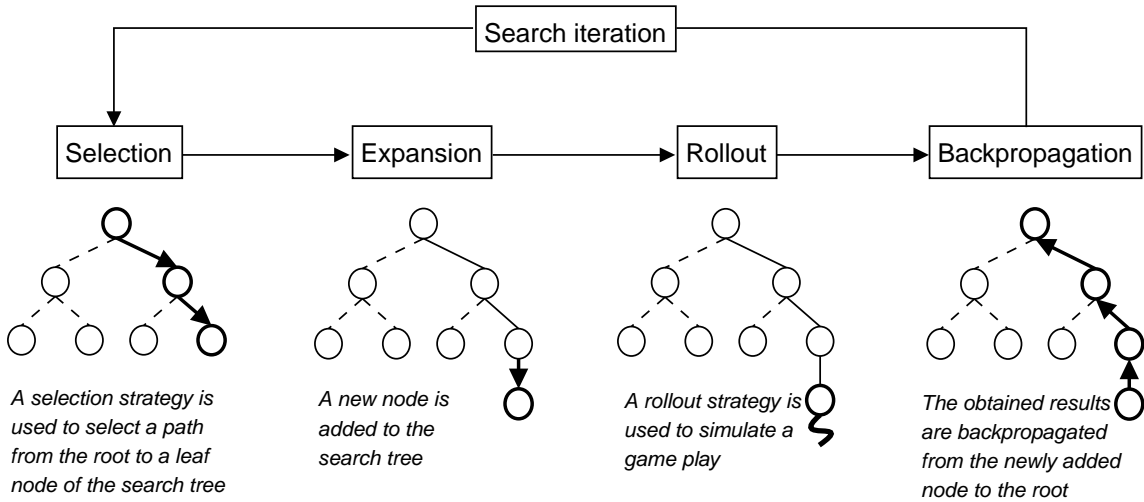
---

A game can easily be modelled as an *multi-armed bandit problem* by representing the different moves of a player as arms. Each time an arm is pulled, a reward is drawn through a simulation exactly the same way as in Monte Carlo evaluation rollout, i.e. players play randomly from the corresponding state until a terminal state is reached. The obtained utility of the player is then used as reward. After a certain number of iterations, the bandit algorithm returns the move with the highest average reward.

## 3.7 Monte Carlo Tree Search

Monte Carlo Tree Search (MCTS) is an algorithm which combines tree-search and Monte Carlo evaluations without separating the min-max phase (or  $max^n$  if the game isn't a two player zero-sum game) from the Monte Carlo phase [Coulom, 2006]. More precisely, MCTS is a breadth-first search method which builds iteratively a search tree and runs Monte-Carlo rollouts at its leaf nodes. Each node represents a possible future state of the game and diverse statistics are stored such as the average score found in the corresponding subtree or the number of visits. The directed links between nodes and their children correspond to the actions leading to the next states. MCTS repeats a process named search iteration generally divided into four phases until some predefined computational budget (time, memory, iteration constraint) is reached. (1) The *selection phase* selects a path from the root to a leaf node of the search tree. (2) The *expansion phase* chooses one or more children to be added to the search tree from the selected leaf node according to the available actions (3) The *rollout phase* simulates the outcome of the game from the newly added node. (4) The *backpropagation phase* propagates backwards the outcome of

the game from the newly added node to the root in order to update the diverse statistics stored in each selected node. The different steps of MCTS are displayed in Figure 3.7.



**Fig. 3.7.:** Monte Carlo Tree Search scheme

### 3.7.1 Selection

The selection phase consists in selecting a path from the root to a leaf node of the search tree. A leaf node of the search tree is either a terminal node of the game or a node which has not got all of its children yet included in the search tree. The choice of the selected child throughout the path is done by using a selection strategy which balances exploration and exploitation. As the selection of a child can be seen as a *multi-arm bandit problem*, bandit algorithms are often used as the selection strategy. The most popular bandit algorithm used for MCTS is the *Upper Confidence bounds applied to Trees* (UCT) proposed by Kocsis and Szepesvári [Kocsis, 2006]. It is a direct application of the UCB algorithm seen in Subsection 3.6. From a selected node  $y$ , the selection strategy UCT chooses the child  $x$  with the highest score  $q_x$ :

$$q_x = \frac{r_x}{n_x} + C \sqrt{\frac{\log(n_y)}{n_x}} \quad (3.4)$$

where  $r_x$  is the sum of rewards found in the subtree with root  $x$ ,  $n_x$  is the number of visits of child node  $x$ ,  $n_y$  is the number of visits of parent node  $y$  and  $C$  the exploration factor.

Other bandit-based selection strategies have also been used in the literature such as *UCB1-Tuned* [Auer, 2002a] or *Exp3* [Auer, 1995].

Many tree-policy enhancements have been proposed throughout the years [Browne, 2012; Świechowski, 2023]. The most famous are probably *All moves as first* (AMAF) [Gelly, 2011] and *Rapid action value estimation* (RAVE) [Gelly, 2007] initially created for the game Go. Both methods aim to share knowledge between related nodes in the search tree. The major assumption behind these methods is that a move will generate similar value if played later on. Hence, one can share statistics between actions played at different states. This results in fast but biased estimates of the action values.

In this thesis, we especially focus on methods which incorporate domain knowledge into the selection strategy. For instance, one could add small rewards or penalties to moves matching some criteria. This biases the selection strategy in favour of moves considered as good according to some expertise. In the game Catan, Szita et al. [Szita, 2009] considerably increase the performance of their MCTS by adding virtual wins to their selection strategy for settlement-building actions and city-building actions.

### 3.7.2 Expansion

The expansion phase consists in choosing which children of the leaf node obtained in the selection phase are expanded to the search tree. One or more children can be added at each expansion step. A common expansion strategy is to select randomly only one non-expanded child of the selected leaf node and add it to the search tree [Coulom, 2006].

Choosing wisely which node to add to the search tree can significantly enhance an MCTS performance. Indeed, expanding nodes related to favourable actions first can be source of improvement. This is generally performed by using a move ordering function. Moreover, in games with a high branching factor, the search tree grows sideways which prevents in-depth inspection of promising branches. The aim of *action reduction* methods is to reduce this effect by adding a limited number of children for each node. For instance, one could prune non-expanded nodes which are related to bad actions according to some expertise or heuristic knowledge. This can enhance an MCTS performance as it reduces the search space and avoids useless exploration. In the game Diplomacy, expanding nodes using the most promising actions first and pruning non-expanded nodes which don't reach a minimal value according to a certain metric considerably improves the quality of the MCTS used for the strategic component of this game [Theodoridis, 2020]. Furthermore, an interesting approach considering tight time constraints is *progressive unpruning* [Chaslot, 2008a]. At first, only a limited number of children are added to the search tree. For instance, in Mango (a famous MCTS applied to Go), only a maximum of 5 nodes are initially expanded for

each parent node in the search tree [Chaslot, 2008a]. Then, each time the number of visits of a parent node reaches a certain level, new children are added to the search tree. By progressively unpruning children, one can obtain a good compromise between in-depth inspection and exploration given tight time constraints.

### 3.7.3 Rollout

In the rollout phase, moves are played starting from the newly added node using a rollout strategy until the game ends in order to simulate the outcome of the game from this particular node. In  $n$ -player games, the outcome is a vector of size  $n$  where each index corresponds to the utility obtained by each player at the end of the payout. As it was the case for the minimax algorithm, in a two-player zero sum game, a single value is sufficient.

To simulate a game, the default rollout strategy used in MCTS is usually to play randomly. Applying search techniques or incorporating domain knowledge in rollout strategies can significantly improve the quality of the simulations by making them more realistic and, hence, improve the overall performance of the algorithm. For instance, using a minimax algorithm at search depth 2 as rollout strategy considerably increases the win rate of MCTS in the game Connect-4 [Baier, 2014]. More specifically, this minimax algorithm could find forced wins and avoid forced losses by seeing two steps ahead. If neither were found, it returns a random move. Domain knowledge can also be easily integrated in the rollout strategy through patterns. More precisely, if a pattern is detected by a player during the rollout phase, then it will play a specific move. For example, MoHex [Arneson, 2010] uses bridge patterns in its rollout strategy in the game Hex.

Given a certain time budget, using too sophisticated rollout strategies may be unproductive. Indeed, if computing each move in the rollout is too expensive, this will drastically diminish the number of search iterations performed by the MCTS and, hence, impact its performance. Thus, it is important to achieve good sampling by computing realistic playouts without consuming too much time. Moreover, if the rollout strategy is too deterministic, the exploration of the search space will be too selective and, hence, the quality of the sampling will suffer [Szita, 2009].

Another issue which is worth pointing out is that, when a player needs to play an action quickly, running a rollout until the game ends might not be very effective, especially if the terminal nodes of the corresponding game have great depth. One possible alternative is to terminate the rollout early at an arbitrary depth and apply an evaluation function to the

last state encountered. It has been shown in [Lorentz, 2016] that even a weak evaluation function can sometimes outperform long random rollouts in certain board games.

### 3.7.4 Backpropagation

The backpropagation phase consists in propagating backwards the outcome of the game obtained in the rollout phase from the newly added node to the root node to update the statistics stored in each selected node. The most popular backpropagation strategy is to save, for each node of the search tree, the average utilities of all rollouts performed through this node.

## 3.8 Related works on MCTS

In this subsection, we give an overview of previous applications of MCTS in many classes of games. We especially focus on games which share similar general game properties with our bidding game in SAA.

### 3.8.1 Application of MCTS to two-player deterministic zero-sum games with perfect and complete information

As Artificial Intelligence (AI) research on games has mainly focused on two-player zero-sum deterministic games with discrete action spaces, perfect and complete information, MCTS has especially been applied to such games [Browne, 2012]. This tree-search algorithm is mostly known for its application to connection games such as *Hex* [Arneson, 2010] or *Havannah* [Teytaud, 2009] and combinatorial games such as *Go* [Lee, 2009; Coulom, 2006] or *Othello* [Robles, 2011] where it achieved state-of-the-art performance. As already mentioned in Subsection 3.7.1, the most popular variant of MCTS is UCT. Its popularity is especially due to the contributions of Kocsis and Szepesvári [Kocsis, 2006] who proved that the probability of playing a suboptimal action with UCT converges to zero at polynomial rate as the number of search iterations grows to infinity. In other words, the search tree of UCT converges to the minimax tree and, therefore, the selected action is optimal. However, such guarantees only hold for two-player zero sum games with perfect and complete information.



### 3.8.2 Application of MCTS to n-player games

In n-player zero-sum games, a common approach is to use MCTS-max<sup>n</sup> [Nijssen, 2013]. This algorithm uses a max<sup>n</sup> tree structure. Its UCT variant achieves promising results in games such as *Chinese Checkers* [Sturtevant, 2008]. Other variants are proposed in [Nijssen, 2012a] using Paranoid and Best-reply tree structure in n-player games such as *Chinese Checkers*, *Blokus*, *Focus* and *Rolit*. However, using such tree structure seems to lead generally to weaker MCTS algorithms than its max<sup>n</sup> variant.

Little work has been done regarding applications of MCTS to n-player non zero-sum games. This is due to the fact that the majority of adversarial games have only two possible outcomes: either there is one winner and the rest of the players lose or it is a draw.

### 3.8.3 Application of MCTS to non-deterministic games with chance events

In non-deterministic games presenting chance events, two approaches are generally privileged to apply MCTS. The first consists in adding explicit chance nodes to the tree structure of the MCTS [Veness, 2011]. As it was also the case for expectimax search, this is only possible if the probabilities of each decision taken by Nature are known. An alternative approach is to use an open-loop variant of MCTS where the nodes in its search tree no longer represent the different states of the game but a sequence of actions leading to them from the root node. This results in a considerably smaller search tree than if chance nodes were explicitly added as each sequence of actions corresponds to only one path in the search tree. This open loop MCTS approach ended up becoming very successful and even won the 2014 General Video Game Playing Competition (GVG-AI) [Perez-Liebana, 2015]. Given an infinite amount of time, the first approach will obviously perform better as decisions will be taken in the search tree knowing exactly in which state it is while in the second approach the same sequence of actions could lead to completely different states. However, if the thinking time is relatively small, the second approach can perform better than the first as it handles a smaller search tree.

### 3.8.4 Application of MCTS to simultaneous move games

Until now, we have essentially focused on adversarial search methods initially designed for turn-based games. In simultaneous move games, an interesting approach is to use

a Simultaneous Move MCTS (SM-MCTS). The main difference with the standard tree structure of MCTS is that a *joint move* is selected at each node of the search tree in the selection phase [Tak, 2014]. This introduces the indeterminism caused by the opponents in changing the game state while a player takes an action into the tree structure of the MCTS. Thus, the tree structure of SM-MCTS is more representative of the actual game as players select their moves independently from one another. The most popular variant of SM-MCTS uses UCT as selection index and is named *Decoupled UCT* (DUCT). However, DUCT has been shown not to converge to the Nash equilibrium in two-player zero-sum simultaneous move games even in one shot single state games [Schaeffer, 2009]. Nevertheless, DUCT has been successful in general game playing. Another approach that achieves surprisingly good performance in many simultaneous move games is to use the standard Sequential UCT (SUCT) [Schaeffer, 2009]. However, as, in the selection phase, the move selected by the root player is revealed before the end of the turn to its opponents, they can adjust their actions given this extra information. Thus, SUCT tends to play defensively which can be a handicap in some games.

It is known that the optimal policy at a simultaneous move node can be mixed [Cowling, 2012a]. Hence, Teytaud and Flory suggest in [Teytaud, 2011] to use *EXP3* which returns a mixed policy rather than *UCB* at simultaneous move nodes in the selection phase. Furthermore, although MCTS is not explicitly designed to return a mixed policy, it usually does in the sense that it can return different actions when applied to a same state. This comes from the random nature of the rollout phase. Hence, MCTS applied to simultaneous move games are usually not deterministic. For instance, an MCTS applied to the game Rock Paper Scissors will usually return with the same probability one of the three actions.

### 3.8.5 Application of MCTS to imperfect information games

In the literature, we identify three types of MCTS which are applied to imperfect information games: *Determinized MCTS*, *Information-State MCTS* and *Belief-State MCTS*.

- Determinization is one of the most popular approach to compute a good strategy in imperfect information games. It generates an instance of the equivalent deterministic game of perfect information where all hidden and random information is assumed known to all players. For example, a determinization of the game Phantom Go (a variant of Go where a player can not see their opponent's stones) is an instance of the actual game Go starting from a particular state. This state is generated such that it is compatible with the player's observations of the Phantom Go game played

until now. From this state onwards, all stones played are visible. Many determinized games are generated and a standard MCTS is applied to each different instance. The results obtained in each independent search tree are then combined. This is known as separate-tree determinization and is similar to root parallelisation [Chaslot, 2008b]. Different combinations are possible. For example, the *Ensemble UCT* method [Fern, 2011] sums for each pair of state-action the utilities obtained in all the different search trees and divides them by the total number of visits. The action at the root which maximises the preceding score is then selected. Another possible combination is to play the action which has been the most selected across the different search trees. This was used to combine the results of many UCT in the card game Dou Di Zhu [Whitehouse, 2011]. However, determinization faces three main issues:

- It is *memory consuming*. This roughly implies that, for a same fixed memory budget, if we build  $N$  determinized MCTS trees, each search tree must perform  $N$  times less search iterations than if it was built on its own.
  - *Strategy fusion* [Frank, 1998]: In different determinizations, a search algorithm might select different actions in states belonging to a same information set. However, theoretically, a player is supposed to play the same way in states belonging to a same information set. Hence, this constraint imposed on a player's strategy is broken when one combines strategies computed in deterministic games with perfect and complete information.
  - *Non-locality* [Frank, 1998]: This issue is raised by the fact that search algorithms determine the best play at an internal node by only considering the subtree starting from this specific internal node. However, in imperfect information games, values of nodes may be impacted by decisions higher up the tree where opponents steer the game towards certain states and away from others which are indistinguishable for the concerned player (they belong to the same information state) [Cowling, 2012b]. Hence, a move may appear optimal at an internal node by only considering its subtree but not optimal if considering the whole search tree. This issue does not happen in perfect and complete information games.
- To reduce the impact of the two first main issues of determinization, one approach is to consider an MCTS search tree with nodes which no longer represent states, as it was the case for the Determinized MCTS, but information sets. This is known as Information Set Monte Carlo Tree Search (ISMCTS). By using ISMCTS, it is no longer necessary to build many trees for each determinization. This partially solves

the issue of managing one's computational budget. Search iterations are divided into five steps: determinization, selection, expansion, rollout and backpropagation. The determinization phase consists in sampling a state from the current information state and then only use nodes/actions in the search tree compatible with the selected determinization. Some variants of ISMCTS completely eliminate strategy fusion. For instance, Single-Observer ISMCTS with Partially Observable Moves (SO-ISMCTS POM) and Multiple-Observer ISMCTS (MO-ISMCTS) [Cowling, 2012a] both eliminate the strategy fusion issue. However, none of these algorithms is capable of eliminating the non-locality issue. This is mainly due to the fact that these algorithms do not model belief distributions and sample determinizations uniformly at random. In the game with imperfect information *Lord of the Rings: The Confrontation*, Cowling et al. show in [Cowling, 2012a] that ISMCTS methods largely outperforms determinized UCT because they devote their entire computational budget to a single tree.

- In order to reduce the impact of non-locality, a solution is to integrate opponent modelling and calculation of belief distributions to the MCTS. This is the aim of Belief-State MCTS (BS-MCTS) [Wang, 2015]. The nodes of its search tree represent belief-states, i.e. each node represents a vector of states with their respective estimated probability that the game is actually in that state. These probabilities, or beliefs, are learned online using two methods known as *Opponent Guessing* and *Opponent Predicting*. The algorithm has a sampling stage where a state is selected according to the belief distributions. After the sampling stage, all hidden and random information is revealed to the players. This method outperformed Determinized UCT, SO-ISMCTS, MO-ISMCTS in Phantom Go [Cowling, 2012a]. Moreover, in Phantom Go, this method completely eliminates strategy fusion and partially solves non-locality.

It is important to note that we have only focused here on how one could reduce the impact of the three main issues induced by determinization. We insist on the fact that it does not imply that one approach is generally better than the other. Indeed, depending on the game, other issues may also impact each approach differently. For instance, in the game *Phantom (4,4,4)* (a  $m,n,k$ -game with imperfect information), combining many Determinized MCTS largely outperforms SO-ISMCTS and MO-ISMCTS given enough CPU time [Cowling, 2012a].

### 3.8.6 Application of MCTS to games with incomplete information

Methods relying on a determinization process can easily be applied to games with incomplete information as hidden information is then assumed to be known by all players. Hence, the same MCTS methods presented for games with imperfect information can be applied to games with incomplete information. For instance, combining multiple determinized MCTS trees is used in the game *Magic: The Gathering* [Cowling, 2012b]. MCTS has also been applied successfully in Poker by integrating opponent modelling [Van den Broeck, 2009; Ponsen, 2010].

### 3.8.7 Applications of MCTS to auction games

In 2018, Chowdhury et al. combined machine learning methods for the prediction of clearing prices and MCTS to plan a bidding strategy across multiple time periods in Periodic Double Auction [Chowdhury, 2018]. However, with exception of our work, MCTS has never been applied to SAA.

## 3.9 Transposition table

As different sequences of actions can lead to a same game state, different nodes throughout the MCTS search tree may represent the same game state. Reevaluating these nodes each time from scratch and storing their statistics independently can become a major source of inefficiency. Through the use of *transposition tables* [Childs, 2008], one can store information about a game state once and then share it between its corresponding nodes. Thus, any update about the statistics of a node will update the statistics of all nodes representing the same game state. A transposition table is a hash table which maps a game state (or information set) to information related to it. This information can be diverse statistics such as the sum of rewards of a player or the number of times nodes corresponding to this game state has been visited. This information can generally be retrieved in two steps.

The first step consists in converting a game state in a nearly unique integer through a *hash function*. The integer assigned by the hash function to a game state is named *hash value*. The best would be to use an injection between the state space and the set of natural integers as hash function. However, in practise, due to memory constraints, only a limited number of natural integers can be computed. This makes impossible the use of an injection as hash function for games with a huge state space and, hence, there is a risk of collision (type-1 error), i.e. a risk that two game states are assigned the same hash value. Zobrist [Zobrist, 1990] proposes an hash function with an arbitrary small probability of collision. This hash function was primarily made for board games but the idea can easily be extended to other games. Suppose the game can be represented as a board of size  $l$  and the game presents  $w$  different types of pieces. For example, in Chess, the board is of size  $l = 64$  and the game has  $w = 12$  different types of game piece (white pawns, black pawns, white queen, ...). There are  $l \times w$  different couples which can be formed by taking a location on the board and a specific type of game piece. For each of these couples, a pseudo-random number is generated and stored. The Zobrist hash function consists in applying the XOR operator to all the pseudo-random numbers of the couples present in a game state. Depending on the game, these pseudo-random numbers are 32-bit or 64-bit long. If the transposition table is of size  $2^k$ , the  $k$  first bits of the generated hash value are used to find the location of the game state in the transposition table. This is called the *hash index*. The remaining bits are used to differentiate game states which have the same hash index in order to avoid collisions (type-2 error). This is called the *hash key*. The hash key is saved in the entry of the transposition table. The first step consists in computing the index of the transposition table associated to a game state. The second step is just using this index to directly access the desired information stored in the transposition table. Transposition tables considerably reduce the search space and, hence, are a major enhancement for search algorithms. However, they can be the source of two different errors: *type-1 errors* and *type-2 errors*. Type-1 error occurs when two different game states share the same hash value. As the hash indices and hash keys are the same, it is difficult to detect such an error. These two different game states will then share the same entry in the transposition table. This will lead to inaccurate information on each game state, especially in the case of stored statistics. Type-2 error occurs if the hash index are the same for two different game states. However, compared to type-1 error, this can be easily detected as the hash keys are different. Thus, by a simple check of the hash keys, sharing the same information between two different game states can easily be avoided in the case of type-2 errors.

## 3.10 Conclusion

As a multi-round auction can easily be represented as a game in extensive form, using adversarial search methods to compute an efficient bidding strategy in SAA seems to be an appropriate research direction. However, not all aforementioned adversarial search methods are suitable for our problem. Indeed, as SAA is an  $n$ -player game, minimax search and  $\alpha\beta$  pruning are not applicable. Moreover, using search techniques which transform a  $n$ -player game in a minimax game such as Paranoid search or Best-reply search seems meaningless. Indeed, as these search methods assume that the aim of the opponents is to diminish at all cost the root player's utility, the value of the game backtracked at the root node will either be zero or negative as the root player will either end up with no items or exposed. Thus, the root player will be encouraged to drop out of the auction. Hence, the most pertinent tree structure appears to be  $\max^n$  where each opponent tries to maximise their own utility.

SAA admits a huge information set space. It is practically impossible to explore the whole game tree. Thus, as already explained in Subsection 3.5, we can use either heuristic evaluation functions or Monte Carlo evaluation to overcome this issue and assign values to non-terminal nodes of a search tree. As constructing an efficient heuristic evaluation function is very complex, we decide to use Monte Carlo evaluation in this thesis. More precisely, we decide to use MCTS which combines tree-search and Monte Carlo evaluation. The tree structure of the search tree will obviously be  $\max^n$  given what was stated previously. Our aim throughout the thesis is to build an MCTS capable of bidding efficiently in SAA by tackling simultaneously its four main strategical issues: the exposure problem, the own price effect, the budget constraints and the eligibility management problem.

In the following chapters, we are going to introduce different MCTS approaches depending on the general game properties of the bidding game considered as well as on the strategical issues being addressed. All considered bidding games are  $n$ -player non zero-sum games. We sum up these algorithms in Table 3.1.

Introduced	Name	General game properties	Strategical issues	Selection strategy	Small presentation
Chapter 4 (Section 4.4)	$MS^\lambda$	Turn-based Deterministic Perfect information, Complete information	Exposure Own price effect	UCT with selection penalties	Standard MCTS based on two risk-aversion hyperparameters $\lambda^r$ and $\lambda^o$
Chapter 5 (Section 5.4)	$SMS^\alpha$	Simultaneous moves Stochastic Imperfect information, Complete information	Exposure Own price effect Budget constraints Activity rules	UCT with scalarised rewards	SM-MCTS based on risk-aversion hyperparameter $\alpha$
Chapter 6 (Section 6.3)	$SMS_{EXP3}^\alpha$	Simultaneous moves Stochastic Imperfect information, Complete information	Exposure Own price effect Budget constraints Activity rules	EXP3 with scalarised rewards	Replacing UCT in $SMS^\alpha$ by EXP3
Chapter 6 (Section 6.4.1)	$\overline{SMS_{EXP3}^\alpha}$	Simultaneous moves Stochastic Imperfect information, Incomplete information	Exposure Own price effect Budget constraints Activity rules	EXP3 with scalarised rewards	Using expectation as determinization and applying $SMS_{EXP3}^\alpha$ to the resulting game
Chapter 6 (Section 6.4.2)	$DSMS_{EXP3}^\alpha$	Simultaneous moves Stochastic Imperfect information, Incomplete information	Exposure Own price effect Budget constraints Activity rules	EXP3 with scalarised rewards	Generates separate trees for each determinization on which $SMS_{EXP3}^\alpha$ is applied
Chapter 6 (Section 6.4.3)	$SDSMS_{EXP3}^\alpha$	Simultaneous moves Stochastic Imperfect information, Incomplete information	Exposure Own price effect Budget constraints Activity rules	EXP3 with scalarised rewards	Single-tree determinization (as ISMCTS) with the use of a variant of $SMS_{EXP3}^\alpha$

**Tab. 3.1.:** Summary of all MCTS algorithms introduced in this thesis





# Turn-based deterministic Simultaneous Ascending Auction with complete information

## Contents

---

4.1	Deterministic model with complete and perfect information of SAA (d-SAA)	94
4.1.1	Mechanism	94
4.1.2	Bidders' modelling	95
4.1.3	Representation of the d-SAA game in extensive form	95
4.1.4	Comparing d-SAA and SAA	96
4.2	Complexities induced by the d-SAA game	96
4.2.1	General game properties	96
4.2.2	Game complexities	97
4.2.3	Strategical complexities	100
4.3	A new prediction method: frontier prediction of final prices	100
4.3.1	Overestimation of final prices	101
4.3.2	Frontier prediction of final prices	102
4.4	MCTS bidding strategy	107
4.4.1	Selection	107
4.4.2	Expansion	110
4.4.3	Rollout strategy	110
4.4.4	Backpropagation	111
4.4.5	Final move selection	111
4.4.6	Transposition table	111
4.4.7	Interpretation of $MS^\lambda$	112
4.5	Experiments	113
4.5.1	Test cases	114
4.5.2	Extensive experiments	115
4.6	Conclusion	120

---

In this chapter, we present the first efficient algorithm which tackles simultaneously the exposure problem and the own price effect. It is based on an MCTS which uses a new prediction method of final prices for its rollout strategy.

We start by introducing a simplified model of SAA which is turn-based, deterministic, with complete and perfect information. This model does not present any activity rules and bidders are assumed to have unlimited budgets. This enables us to focus only on two of the four main strategical issues in SAA. We name this simplified version *d-SAA*. We then present the different complexities induced by the bidding game in *d-SAA*. We then introduce a new concept for the prediction of final prices named *frontier prediction of final prices*. We show that, in simple environments, the frontier prediction of final prices can be computed in *d-SAA* through the convergence of a specific sequence. We then present our bidding strategy based on MCTS. We name it  $MS^\lambda$ . Its rollout strategy relies on our new prediction method of final prices. Finally, we show through extensive numerical experiments on typical examples of the literature and a large number of random instances that  $MS^\lambda$  significantly outperforms state-of-the-art algorithms in SAA presented in Section 2.4 and achieves higher expected utility by taking less risks.

Most of the results presented in this chapter have been published in our article *Monte Carlo Tree Search Bidding Strategy for Simultaneous Ascending Auctions* [Pacaud, 2022].

## 4.1 Deterministic model with complete and perfect information of SAA (*d-SAA*)

### 4.1.1 Mechanism

In this chapter, we focus on a deterministic turn-based model with perfect information of SAA [Milgrom, 2000; Cramton, 2006; Wellman, 2008] which we refer to as *d-SAA*. In this specific auction,  $m$  indivisible goods are sold via separate and concurrent English auctions between  $n$  players. Bidding occurs in multiple rounds. Players take turns bidding such that, at each round, only the designated player is allowed to bid. The *bid price* of item  $j$ , denoted by  $P_j$ , corresponds to the highest bid obtained so far for item  $j$  and, thus, is its current selling price. At the beginning of each round, the bid price and the current winner of each item is announced to each player. Hence, players are aware of all past actions

taken by their opponents and the induced bidding game is with perfect information. All bid prices are initialised to zero. As in the original SAA [Milgrom, 2000], the auction ends when none of the players have submitted an admissible bid during their last round. This is equivalent in our model to saying that the bid price of each item has not increased during  $n$  consecutive rounds. After the auction closes, the items are sold at their current bid price to their corresponding temporary winners. No activity rules are considered in this chapter.

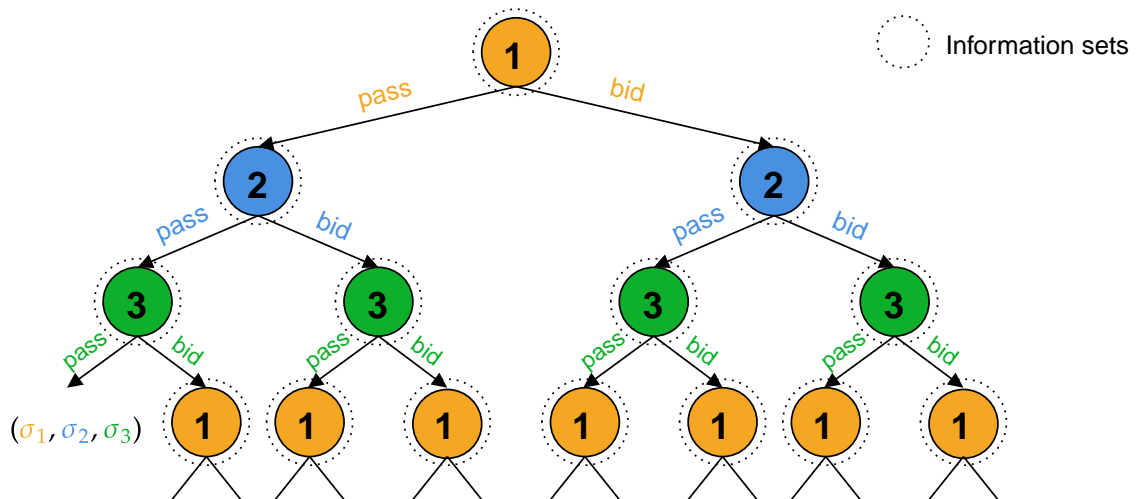
### 4.1.2 Bidders' modelling

As we do not consider activity rules or budget constraints in this chapter, a bidder  $i$  is only defined by its value function  $v_i$  and its utility function  $\sigma_i$ . A first reasonable assumption is to consider both features as common knowledge [Szentes, 2003a; Szentes, 2003b]. Hence, the bidding game induced by d-SAA is a complete information game. This assumption will be relaxed in Chapter 6.

Moreover, in our model, a bidder can only submit on item  $j$  a bid  $P_j^{new}$  equal to its current bid price  $P_j$  plus the fixed bid increment  $\varepsilon$ :  $P_j^{new} = P_j + \varepsilon$ . This simplification of the bidding space is common in the literature on ascending auctions [Goeree, 2014; Wellman, 2008]. We make the classical assumption that a bidder will not continue bidding on an item that it has temporarily won [Wellman, 2008]. Hence, in our model, a winner will always pay a price for an item at most  $\varepsilon$  above the highest opponent bid.

### 4.1.3 Representation of the d-SAA game in extensive form

As many sequential turn-based games, the bidding game induced by d-SAA can be represented in extensive form. Each node corresponds to a different state of the game and each edge represents a different feasible bid. In a d-SAA game, a state is defined by three features: the next player to bid, the bid price and the temporary winner of each item. These three features are common knowledge. As a d-SAA game is a game with perfect information, each player knows exactly which moves have been taken in the past by its opponents and, thus, each information set is a singleton. In other words, an information set corresponds to only one state. An example of a d-SAA game in extensive form with 3 players and one item (a player can either bid or pass) is displayed in Figure 4.1 with its information sets. In this example, when three players pass consecutively, the auction closes and each bidder  $i$  obtains its respective utility  $\sigma_i$ .



**Fig. 4.1.:** A d-SAA game in extensive form with three players and one item

#### 4.1.4 Comparing d-SAA and SAA

The only difference between the mechanism used in d-SAA and the one originally used in SAA presented in Section 1.3.2.1 is that players take turns bidding. By this change, we eliminate stochasticity and simultaneity from our problem. Indeed, ties between players having bid the same amount on a given item can not occur in d-SAA. Hence, the temporary winner is no longer selected randomly amongst these players. This facilitates future studies and the conception of a simpler tree-search algorithm with no chance nodes [Veness, 2011] and with a closed-loop implementation [Perez Liebana, 2015]. From the bidder's point of view, both mechanisms are strategically equivalent if the bid increment is small.

## 4.2 Complexities induced by the d-SAA game

### 4.2.1 General game properties

The game induced by d-SAA has the following general game properties:

- n-player game
- Non-zero sum game

- Sequential: Players take turns bidding.
- Deterministic: As ties do not occur, an action played by a bidder always leads to the same next state.
- Perfect information
- Complete information

## 4.2.2 Game complexities

We compute below the information set space complexity and the game tree complexity of a d-SAA game. It is important to recall that, as the bidding game is with perfect information, the information set space complexity is equal to the state space complexity (see Section 2.2.2). With the assumption of finite valuations, the auction finishes after a finite number of rounds. Thus, we restrict our analysis on both complexities to a given number of rounds  $R$ .

**Theorem 4.1.** *Let  $\Gamma$  be an instance of the d-SAA game. Let  $n$ ,  $m$  and  $R$  be respectively the number of players, the number of items and the number of rounds. The state space complexity of  $\Gamma$  is:*

$$\sum_{i'=0}^{n-1} \left(1 + \sum_{i=0}^{n-1} (R - i - i')^+ \right)^m \mathbb{1}_{\{R \geq i'\}} \quad (4.1)$$

where  $\mathbb{1}_{\{z \geq w\}}$  is the indicator function, equal to 1 if  $z \geq w$  and 0 otherwise.

*Proof.* Let  $\Gamma$  be an instance of a d-SAA game. Let  $n$ ,  $m$  and  $R$  be respectively the number of players, the number of objects and the number of rounds in  $\Gamma$ .

In d-SAA, the index of each player corresponds to the first round where the player is given the opportunity to bid. For example, player  $i$  can only submit its first bid at round  $i$ . Without loss of generality, let's reindex the players such that player 1 corresponds to the player bidding at round  $R$ , player 2 corresponds to the player bidding at round  $R - 1$  and so on.

Let's calculate the number of different states that can take an item  $j$  given that the next bidder to bid is bidder  $i' = 1$ . After  $R$  rounds, item  $j$  remains either unsold or a bidder  $i$  acquires it. A bidder  $i$  can acquire item  $j$  at different prices depending on its position in

the game. Indeed, if  $i < R$ , the last opportunity to submit a bid for player  $i$  is at round  $R - i + 1$ . Thus, the number of different prices that can take item  $j$  is  $R - i + 1$  as the price of an item can only be raised by  $\varepsilon$  each round. If  $i \geq R$ , then player  $i$  will never have the opportunity to bid and acquire item  $j$ . Thus, given that the next bidder to bid is bidder 1, the number of different states that item  $j$  can take if won by bidder  $i$  in  $\Gamma$  is  $(R - i + 1)^+$  with  $z^+ = \max(0, z)$ . The number of different states that can take any item  $j$ , given that the next bidder to bid is bidder 1, is then:

$$\sum_{i=1}^n (R - i + 1)^+ + 1 \quad (4.2)$$

As all items are mutually independent, the number of different states given that the next bidder to bid is bidder 1 is:

$$\left( \sum_{i=1}^n (R - i + 1)^+ + 1 \right)^m \quad (4.3)$$

The number of different states given that the next bidder to bid is  $i'$  can easily be computed by noticing that the last opportunity to submit a bid for bidder  $i'$  is  $R - i' + 1$ . Thus, we apply the same reasoning as above for  $R - i' + 1$  rounds. This gives us that the number of different states given that the next bidder to bid is bidder  $i'$  and  $R - i' + 1 \geq 0$  is:

$$\left( \sum_{i=1}^n (R - i - i' + 2)^+ + 1 \right)^m \quad (4.4)$$

Thus, the state space complexity of  $\Gamma$  is:

$$\sum_{i'=1}^n \left( \sum_{i=1}^n (R - i - i' + 2)^+ + 1 \right)^m \mathbb{1}_{\{R - i' + 1 \geq 0\}} \quad (4.5)$$

where  $\mathbb{1}_{\{z > w\}}$  is the indicator function, equal to 1 if  $z > w$  and 0 otherwise. This can be rewritten as:

$$\sum_{i'=0}^{n-1} \left( 1 + \sum_{i=0}^{n-1} (R - i - i')^+ \right)^m \mathbb{1}_{\{R \geq i'\}} \quad (4.6)$$

□

**Theorem 4.2.** *Let  $\Gamma$  be an instance of the  $d$ -SAA game. Let  $n$ ,  $m$  and  $R$  be respectively the number of players, the number of items and the number of rounds in  $\Gamma$ . A lower bound of the game tree complexity of  $\Gamma$  is:*

$$2^{m(n-1) \lfloor \frac{R}{n} \rfloor} \quad (4.7)$$

*Proof.* It is important to remember that only one player bids per round of the d-SAA game whereas all players bid during a round of the SAA game. To compute a lower bound of the game tree complexity of the d-SAA game, we only need to notice that more branches are created in  $n$  rounds of the d-SAA game than in one round of the SAA game with a deterministic tie-breaking rule and with the assumption that no players bid on temporary won items. Each time a player bids on an item in the d-SAA game, it directly becomes the temporary winner of this item and, hence, increases the number of bidding possibilities of its opponents. However, in the SAA game, as bids are simultaneous, the bid submission of one doesn't impact the bidding possibilities of the others. Hence, the number of bidding scenarios is inferior in one round of the SAA game with deterministic tie-breaking rule than in  $n$  rounds of the d-SAA game. Let's compute a lower bound of the game tree complexity of the SAA game with deterministic tie-breaking rule.

Let  $\Gamma'$  be an instance of a SAA game with deterministic tie-breaking rule. Let  $n$ ,  $m$ ,  $R$  and  $k$  be respectively the number of players, the number of objects, the number of rounds and the number of admissible bids per object in  $\Gamma'$ . There are  $k + 1$  ways of bidding on a temporary lost item as each player has the choice between not bidding on it or submitting an admissible bid.

Let's first calculate a lower bound of the number of different branches created in the game tree of  $\Gamma'$  during a round.

Suppose player  $i$  is the temporary winner of  $m_i$  objects. Thus, player  $i$  has  $(k + 1)^{m - m_i}$  different ways of bidding. As this is the case for all players, there are  $(k + 1)^{n \times m - \sum_{i=1}^n m_i}$  different bidding scenarios and, hence,  $(k + 1)^{n \times m - \sum_{i=1}^n m_i} - 1$  new branches. Moreover, as  $\sum_{i=1}^n m_i \leq m$ , the number of different branches created during a round is lower bounded by  $(k + 1)^{m(n-1)} - 1$ .

From this, a lower bound of the number of possible different paths can easily be calculated by induction. Indeed, any non-terminal node starting a bidding round of the game tree of  $\Gamma'$  induces at least  $(k + 1)^{m(n-1)} - 1$  new branches during a round and, hence, the game tree complexity of  $\Gamma'$  is lower bounded by:

$$\sum_{l=0}^R ((k + 1)^{m(n-1)} - 1)^l = \frac{((k + 1)^{m(n-1)} - 1)^{R+1} - 1}{(k + 1)^{m(n-1)}} \quad (4.8)$$

Thus, a lower bound of the game tree complexity of a SAA game with deterministic tie-breaking rule  $\Gamma'$  is  $O((k + 1)^{m(n-1)R})$ .



Hence, as  $k = 1$  for a d-SAA game, a lower bound of the game tree complexity of d-SAA game  $\Gamma$  is  $O(2^{m(n-1)} \lfloor \frac{R}{n} \rfloor)$ .

□

*Remark.* The state space and game tree complexity grow respectively polynomially and exponentially with the number of rounds  $R$  for a d-SAA game. In the case of homogeneous environments (when all items are identical), both complexities are considerably reduced as the number of different allocations per player is no longer  $2^m$  but  $m + 1$ . This simplification is often used in the literature [Goeree, 2014; Wellman, 2008].

*Example.* The SAA, which took place in Italy in 2018, had 12 5G spectrum licences sold between 5 telecommunication companies after 171 rounds [European 5G Observatory, 2018]. For such values, the state space and game tree complexities of the corresponding d-SAA game are respectively greater than  $10^{35}$  and  $10^{491}$ .

### 4.2.3 Strategical complexities

In addition to the difficulties generated by the high state space and game tree complexities of the d-SAA game, we focus on its two main strategical issues: the exposure problem and the own price effect. Both issues have already been fully detailed and illustrated in Section 2.2.3. As d-SAA does not present any activity rule and bidders have unlimited budgets, the eligibility and budget management problems are not covered in this chapter.

## 4.3 A new prediction method: frontier prediction of final prices

There is actually a gap in literature regarding the prediction of closing prices of an auction as no current methods meet the three following desirable properties:

1. The concept exists in all auctions
2. It takes into account auctions' particularities

3. The final prediction is independent of a single specific strategy profile.

For example, concepts such as Walrasian price equilibrium [Arrow, 1971] or self-confirming price prediction [Wellman, 2008] don't always exist when preferences exhibit complementarities as it is the case in *Example 1* (see Section 2.2.3). Standard tâtonnement processes, such as the one used to compute the *expected price equilibrium* described in Section 2.4.2.2, return the same price vector regardless of the auction's specificities (e.g., bid increment  $\varepsilon$ ). The final prediction is then completely independent of the auction mechanism of d-SAA which is problematic. Computing an initial prediction by using only the outcomes of a single strategy profile is relevant only if bidders actually play according to this strategy profile. For instance, simulating SAA-c games where all bidders play SB and using the average closing prices as initial prediction is relevant if the actual bidders play SB.

Thus, we introduce in this section a new concept for the prediction of final prices in d-SAA named *frontier prediction of final prices* which meets the three above criteria. This concept relies on the bidding strategy PP which is developed in Section 2.4.2. It is based on the set of overestimation of final prices which we introduce for the first time in this thesis. We then present a simple method based on a specific sequence which converges to the frontier prediction of final prices. We prove this convergence in simple environments. We conjecture that this result stands in more complex environments.

### 4.3.1 Overestimation of final prices

**Definition 4.1.** Let  $\Gamma$  be an instance of the d-SAA game. The function  $f_\Gamma : \mathbb{R}_+^m \mapsto \varepsilon\mathbb{N}^m$  is defined as the final prices obtained in  $\Gamma$  when all players play PP with initial prediction  $p \in \mathbb{R}_+^m$ .

*Remark.* Due to the discrete nature of bidding in an instance  $\Gamma$  of the d-SAA game because of the price increment  $\varepsilon$ ,  $f_\Gamma$  is a deterministic piece-wise constant bounded function. Each region where  $f_\Gamma$  is constant is delimited by a set of linear inequalities on the initial prediction  $p$  of the form:  $e \cdot p \leq b$  with  $b \in \mathbb{R}$ ,  $p \in \mathbb{R}_+^m$ ,  $e \in \{-1, 0, 1\}^m$ . Thus, each region is convex as they are defined as an intersection of hyperplanes.

*Remark.* Theoretically, a closed-form expression of  $f_\Gamma$  can be computed for any instance  $\Gamma$  of the d-SAA game. However, in practice, this can only be done for small instances. Thus,  $f_\Gamma(p)$  is usually computed numerically by simulating a d-SAA where all players play PP

with initial prediction  $p$ . In *Example 1* (see Section 2.2.3), for  $p = (11, 11)$ ,  $f_\Gamma(p) = (1, 0)$  as player 1 bids on the item of lowest index and player 2 does not bid as it predicts that the price for both items is 22.

**Definition 4.2.** Let  $\Gamma$  be an instance of the d-SAA game. A price  $p$  is said to be an overestimation of final prices in  $\Gamma$  if  $p \geq f_\Gamma(p)$  component-wise. Let  $O_\Gamma$  be the set of overestimated final prices in  $\Gamma$ , i.e.,  $O_\Gamma = \{p \in \mathbb{R}_+^m, p \geq f_\Gamma(p)\}$ . Conversely, a price  $p$  is said to be an underestimation of final prices in  $\Gamma$  if  $p \leq f_\Gamma(p)$  component-wise.

*Remark.* If price  $p$  is an underestimation of final prices in game  $\Gamma$  and all players play PP with initial prediction  $p$ , then there exists a round in  $\Gamma$  where, from this point, all players play straightforwardly. The special case where  $p = f_\Gamma(p)$  is referred to as a self-confirming price prediction. Note that under certain conditions, the price vectors belonging to Walrasian equilibria are included in the set of overestimated final prices. To be more precise, a sufficient condition is that  $p_j$  is greater than  $\varepsilon$  for all items and that the demand set  $\operatorname{argmax}_X \sigma(X, p)$  of each player is a singleton.

*Remark.* It is easy to exhibit an instance  $\Gamma$  of a d-SAA game such that  $O_\Gamma$  isn't a closed set. Therefore, for practical issues, we consider the closure  $\bar{O}_\Gamma$  of  $O_\Gamma$ .

### 4.3.2 Frontier prediction of final prices

It is based on a simple approach which uses the smallest overestimated final price as a prediction for any d-SAA game instance  $\Gamma$ . Such a prediction has the merit of existing in all auctions. Thus, the property (1) presented in introduction of this section is fulfilled.

However, the smallest overestimated final price may not be unique and, if so, we should select the ones which are the most relevant to the instance  $\Gamma$  specificities. An effective procedure would be to select the ones which are closest in average to their initial prediction when applied to  $f_\Gamma$  with a slight underestimation. A slight underestimation of a smallest overestimated final price gives relevant information through  $f_\Gamma$  on the strength of the bond that ties the initial prediction to  $\Gamma$ . Indeed, players with high values will participate actively in the auction and the resulting closing prices of some goods will exceed their initial prediction by definition.

However, slightly overestimating a smallest overestimated final price will generally give no information through  $f_\Gamma$  on whether the prediction and  $\Gamma$  are closely related. Indeed, nearly all players will drop out of the auction instantly and, hence, the resulting closing price of each good will rarely exceed  $\varepsilon$ . This is also generally the case without any modification of the initial prediction if the smallest overestimated price isn't a self-confirming price prediction. The smallest overestimated prices which are the most relevant regarding  $\Gamma$  specificities are referred to as the *frontier prediction of final prices*. Thus, property (2) is fulfilled.

As, by construction, our approach doesn't rely on the results of a single specific strategy profile, property (3) is fulfilled. Thus, the frontier prediction of final prices verifies the three above requirements.

**Definition 4.3.** Let  $\Gamma$  be an instance of the d-SAA game. The set of smallest overestimated final prices  $O_\Gamma^{min}$  is defined as  $O_\Gamma^{min} = \operatorname{argmin}_{p \in \bar{O}_\Gamma} \|p\|_1$ .

*Remark.* We choose the Manhattan distance given the fact that the closed-form expression of  $f_\Gamma$  is defined by a set of linear inequalities on its initial prediction.

*Remark.* For any instance  $\Gamma$  of the d-SAA game,  $O_\Gamma^{min}$  is well defined as the closed set  $\bar{O}_\Gamma$  contains the vector of prices  $\max_{i=1, \dots, n} v_i(\{1, \dots, m\}) \times (1, \dots, 1)$  and  $\forall p \in \bar{O}_\Gamma, \forall j \in \{1, \dots, m\}, p_j \geq 0$ .

**Definition 4.4.** Let  $\Gamma$  be an instance of the d-SAA game. We define

$$g_\Gamma : \mathbb{R}_+^m \rightarrow \mathbb{R}_+$$

$$p \mapsto \lim_{\eta \rightarrow 0^+} \frac{1}{\eta^m} \int_{\Omega=[0, \eta]^m} \|f_\Gamma(p - z) - p\|_2^2 dz. \quad (4.9)$$

For any  $\Gamma$ , we conjecture that the lower  $g_\Gamma(p)$  is, the more the prediction of closing prices  $p$  is relevant regarding  $\Gamma$  specificities.

*Remark.* It is important to note that, as  $f_\Gamma$  is a piece-wise constant function,  $g_\Gamma$  is a piece-wise continuous function. Thus,  $g_\Gamma$  doesn't necessarily admit a minimum on  $\mathbb{R}_+^m$ .

**Definition 4.5.** Let  $\Gamma$  be an instance of the d-SAA game. Let  $l_\Gamma^* = \inf\{g_\Gamma(p) | p \in O_\Gamma^{min}\}$  be the lower-bound of  $g_\Gamma$  on the set of smallest overestimated final prices. Let  $\mathbf{B}(q, r)$  be

the ball of radius  $r$  and centre  $q$ . For all  $\gamma > 0$ , we define  $O_{\Gamma,\gamma}^{min,*}$  as the set of smallest overestimated final prices which are  $\gamma$  close to the infimum.

$$O_{\Gamma,\gamma}^{min,*} = \{p \in O_{\Gamma}^{min}, \inf \{g_{\Gamma}(q) | q \in \mathbf{B}(p, \gamma) \cap O_{\Gamma}^{min}\} = l_{\Gamma}^*\} \quad (4.10)$$

**Definition 4.6.** Let  $\Gamma$  be an instance of the d-SAA game. The **frontier prediction of final prices** is defined as:

$$O_{\Gamma}^{min*} = \bigcap_{\gamma > 0} O_{\Gamma,\gamma}^{min*} \quad (4.11)$$

The frontier prediction of final prices is the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_{\Gamma}|_{O_{\Gamma}^{min}}$ , the restriction of  $g_{\Gamma}$  to  $O_{\Gamma}^{min}$ , for any  $\gamma > 0$ .

*Remark.* If  $g_{\Gamma}$  admits a minimum on  $O_{\Gamma}^{min}$ , then  $O_{\Gamma}^{min*} \supseteq \operatorname{argmin}_{p \in O_{\Gamma}^{min}} g_{\Gamma}(p)$ .

*Example.* In *Example 1* (see Section 2.2.3), it is easy to show that the set of overestimated prices is  $O_{\Gamma} = \{p \in \mathbb{R}_+^2, p_1 + p_2 \geq 20\}$ . Thus, the set of smallest overestimated prices is  $O_{\Gamma}^{min} = \{(\alpha, 20 - \alpha), \alpha \in [0, 20]\}$ . Given  $f_{\Gamma}$  (see Appendix A), we have 7 different cases to evaluate depending on the values of  $\alpha$ :

- For  $\alpha \in [0, 8[$ ,  $g_{\Gamma}((\alpha, 20 - \alpha)) = (12 - \alpha)^2 + (20 - \alpha - 1)^2$  so  
 $\inf\{g_{\Gamma}((\alpha, 20 - \alpha)) | \alpha \in [0, 8[ ] = 137$ .
- For  $\alpha = 8$ ,  $g_{\Gamma}((8, 12)) = \frac{1}{2}(4^2 + 11^2) + \frac{1}{2}(4^2 + 1^2) = 77$
- For  $\alpha \in ]8, 10[$ ,  $g_{\Gamma}((\alpha, 20 - \alpha)) = (12 - \alpha)^2 + (20 - \alpha - 11)^2 > 5$  and  
 $\lim_{\alpha \rightarrow 8^+} g_{\Gamma}((\alpha, 20 - \alpha)) = 17$
- For  $\alpha = 10$ ,  $g_{\Gamma}((10, 10)) = \frac{1}{2}((12 - 10)^2 + (11 - 10)^2) + \frac{1}{2}((11 - 10)^2 + (12 - 10)^2) = 5$
- For  $\alpha \in ]10, 12[$ ,  $g_{\Gamma}((\alpha, 20 - \alpha)) = (11 - \alpha)^2 + (20 - \alpha - 12)^2 > 5$  and  
 $\lim_{\alpha \rightarrow 12^-} g_{\Gamma}((\alpha, 20 - \alpha)) = 17$
- For  $\alpha = 12$ ,  $g_{\Gamma}((12, 8)) = 77$
- For  $\alpha \in ]12, 20]$ ,  $g_{\Gamma}((\alpha, 20 - \alpha)) = (\alpha - 1)^2 + (20 - \alpha - 12)^2$  so  
 $\inf\{g_{\Gamma}((\alpha, 20 - \alpha)) | \alpha \in ]12, 20]\} = 137$ .

Thus, the **frontier prediction of final prices** for *Example 1* is  $\{(10, 10)\}$ .

*Remark.* Walrasian price equilibrium or self-confirming price predictions don't exist in *Example 1*, see Section 2.4.2.1 and Section 2.4.3.1 [Wellman, 2008].

Even when a Walrasian equilibrium exists, the frontier prediction of final prices seems to be a more accurate prediction of final prices as it takes into account the mechanism of d-SAA. For instance, consider the following simple example.

*Example.* Consider a d-SAA game with three players, one item and a bid increment  $\varepsilon = 1$ . The first player values the item at 4.5, the second player at 4.7 and the third player at 4.9.

In this game, all Walrasian equilibria can be described with a price vector  $p \in [4.7, 4.9]$  and an allocation vector returning the item to the third player. However, the set of overestimated prices is  $O_\Gamma = [4, +\infty[$  and, thus, the frontier prediction of final prices is the singleton  $\{4\}$ . If the first player decides to play PP using as initial prediction the price vector of a Walrasian equilibrium, then it immediately drops out of the auction. However, if the first player decides to play PP using as initial prediction the frontier prediction of final prices and all other players play SB, then it obtains the item for a price of 4 and has a utility of 0.5. Thus, in this example, using a price vector of the frontier prediction of final prices instead of the Walrasian equilibrium increases one's utility. Moreover, the price returned by the frontier prediction of final prices in this specific example is the unique self-confirming price prediction of the bidding game as  $f_\Gamma(4) = 4$ .

We list hereafter a few properties which have been proven for the frontier prediction of final prices, mostly in simple environments.

**Property 4.1.** *The frontier prediction of final prices exists in any instance  $\Gamma$  of the d-SAA game.*

*Proof.* Existence guaranteed by its construction. □

**Property 4.2.** *Let  $\Gamma$  be an instance of the d-SAA game. If  $m = 2$ ,  $n = 2$ , infinitesimal  $\varepsilon$  and both players have super-additive value functions, then  $O_\Gamma^{min*}$  is a singleton.*

*Proof.* See Appendix B. □

**Property 4.3.** *Let  $\Gamma$  be an instance of the d-SAA game. If  $m = 2$ ,  $n = 2$ , infinitesimal  $\varepsilon$  and both players have super-additive value functions, then the sequence  $P_{t+1} = \frac{1}{t+1}f_{\Gamma}(P_t) + (1 - \frac{1}{t+1})P_t$  with  $P_0$  the null vector of prices converges to the unique element of  $O_{\Gamma}^{min*}$ .*

*Proof.* See Appendix B. □

**Conjecture 4.1.** *Let  $\Gamma$  be an instance of the d-SAA game. The sequence  $P_{t+1} = \frac{1}{t+1}f_{\Gamma}(P_t) + (1 - \frac{1}{t+1})P_t$  with  $P_0$  the null vector of prices converges to an element of  $O_{\Gamma}^{min*}$ .*

The convergence of the sequence  $P_{t+1} = \frac{1}{t+1}f_{\Gamma}(P_t) + (1 - \frac{1}{t+1})P_t$  to the unique element of  $O_{\Gamma}^{min*} = \{(10, 10)\}$  in *Example 1* (see Section 2.2.3) is proved in Appendix A.2 with its convergence rates. Moreover, in our numerical experiments, we observe that sequence  $P_t$  converges in all undertaken d-SAA instances. However, we are unable to tell for complex environments if its limit belongs to the frontier prediction of final prices. In the following remark, we provide a proof direction that we consider promising to show that sequence  $P_t$  converges in all d-SAA instances.

*Remark.* To prove the convergence of sequence  $P_t$  in all d-SAA instances, we believe that one should rely on the same spatial representation for  $f_{\Gamma}$  than the one displayed in Figure A.1. More precisely, as seen in Section 4.3.1, the input space of  $f_{\Gamma}$  can be divided into convex regions for which  $f_{\Gamma}$  is constant (each region is delimited by linear inequalities). Thus, the trajectory of sequence  $P_{t+1} = \frac{1}{t+1}f_{\Gamma}(P_t) + (1 - \frac{1}{t+1})P_t$  will always follow a straight path in these regions. Moreover, it can be shown that sequence  $P_t$  is confined to a bounded space for all iterations  $t$ . From this representation, we think that the general proof can be obtained in two steps: (1) Prove that any sequence with the above spatial representation converges if it does not admit any "cycle" and (2) Prove that, for all d-SAA instances  $\Gamma$ , the spatial representation of  $f_{\Gamma}$  does not admit any cycle. In Figure 4.2, we provide an example of a "cycle". Each coloured region corresponds to a region for which  $f_{\Gamma}$  is constant. For example, in the yellow region,  $f_{\Gamma}(P) = (1, 1)$ . The arrows represent the dynamic of sequence  $P_t$ . In this example, due to the cycle in the spatial representation,  $P_t$  does not converge and goes round the polygon represented in red.

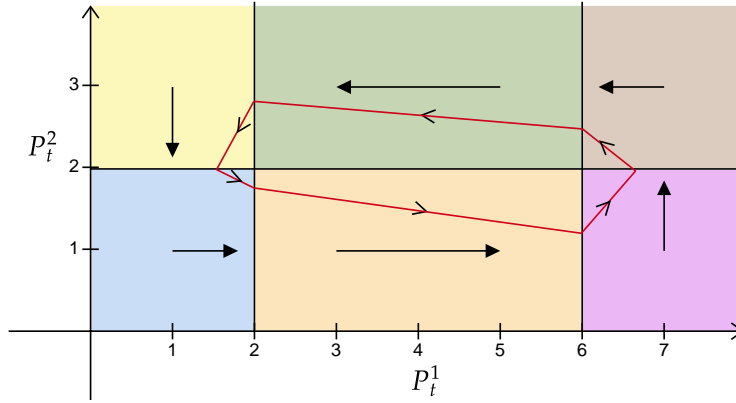


Fig. 4.2.: A fictitious example of a cycle in the spatial representation of  $f_\Gamma$

## 4.4 MCTS bidding strategy

In this section, we present a closed-loop variant of MCTS-max<sup>n</sup> [Nijssen, 2013] which tackles simultaneously the own price effect and the exposure problem. We name it  $MS^\lambda$ . Each node of its search tree  $x$  stores the following statistics: the sum of rewards  $r_x$  obtained when selecting node  $x$ , the number of visits  $n_x$ , the estimated lower bound  $a_x$  and the estimated higher bound  $b_x$  of the reward support. Before running  $MS^\lambda$ , we compute the conjectured limit  $P^* \in O_\Gamma^{min*}$  of sequence  $P_{t+1} = \frac{1}{t+1}f_\Gamma(P_t) + (1 - \frac{1}{t+1})P_t$ . We use this prediction of final prices  $P^*$  for the rollout strategy of  $MS^\lambda$ . The four different phases of a search iteration are described below.

### 4.4.1 Selection

The selection phase consists in selecting a path from the root to a leaf node of the search tree. We propose a new version of selection index UCT [Kocsis, 2006] based on Hoeffding inequality [Hoeffding, 1994] and an online estimation of the size of the reward support. From a selected node  $y$ , the selection strategy chooses the child  $x$  with the highest score  $q_x$ :

$$q_x = \frac{r_x}{n_x} + \max(c_x - a_x, \varepsilon) \sqrt{\frac{2 \log(n_y)}{n_x}} - no\_object(x) - risky\_move(x) \quad (4.12)$$

where  $r_x$  is the sum of rewards obtained when selecting child node  $x$ ,  $n_x$  is the number of visits of child node  $x$ ,  $n_y$  is the number of visits of parent node  $y$ ,  $\varepsilon$  is the bid increment,  $a_x$



is the estimated lower bound and  $c_x$  is the estimated higher bound of the reward support found in the subtree with root  $x$ . We fully detail each term of our selection strategy below:

- The two first terms come directly from the *Hoeffding's Inequality* and give an upper confidence bound on the average reward  $\frac{r_x}{n_x}$  when selecting child node  $x$  using  $\max(c_x - a_x, \varepsilon)$  as an estimation of the size of the reward support.
- The third term  $no\_object(x)$  is a penalty term which has been introduced to avoid players from passing their turn if they have got nothing to lose by bidding on an additional item and, thus, emulate rational behaviour. For instance, if a player is currently winning a set of undesired goods and can increase its utility by directly acquiring another item, it should continue to bid even though its chances of obtaining any other item might be slim. We define  $no\_object(x)$  as the maximum surplus a player could obtain by just acquiring another item under the assumption that the auction ends after its turn. More formally, let  $S_{-j}$  be the set containing all goods except item  $j$ . A player  $i'$  will no longer bid on item  $j$  if  $\forall X \subset S_{-j}, P_j + \varepsilon > v_{i'}(X + \{j\}) - v_{i'}(X)$ . In other words, player  $i'$  will no longer bid on item  $j$  if the maximum surplus it can obtain by winning  $j$  is lower than the minimal admissible bid it needs to submit on  $j$ . We define  $\Pi_j^i = \max_{i' \in \{1, \dots, n\} \setminus \{i\}} \max_{X \subset S_{-j}} v_{i'}(X + \{j\}) - v_{i'}(X) - \varepsilon$  as the minimal price from which item  $j$  is considered as undesired by all opponents of  $i$ . If  $P^x$  is the price vector at child node  $x$ ,  $i$  is the player bidding at parent node  $y$  and  $X_x^i$  the set of goods temporarily won by player  $i$  at  $x$ , the penalty term is defined as

$$no\_object(x) = \begin{cases} \max_{j \in \{1, \dots, m\} \setminus X_x^i} (v_i(X_x^i + \{j\}) - v_i(X_x^i) - P_j - \varepsilon)^+ & \text{if } \{j' \in X_x^i, P_x^{j'} \leq \Pi_{j'}^i\} = \emptyset \\ 0 & \text{otherwise} \end{cases} \quad (4.13)$$

The penalty  $no\_object(x)$  reduces the risk of exposure of the root player. To illustrate this, suppose that player 1 is currently winning both items at vector price  $P = (5, 5)$  in *Example 1* (see Section 2.2.3). It is player 2's turn to bid. Player 2 knows that player 1 will never bid on an item which bid price is strictly greater than 11. Hence, player 2 knows, that by bidding aggressively on both items, it will acquire them in the worst-case scenario for a price of 24. If player 2 decides to bid on both items, then player 1 has the choice between passing, bidding on only one of the two items or bidding on both items. Bidding on both items is unreasonable for player 1 as, if player 2 passes, player 1 ends up with a utility of -2. Thus, the remaining choices for player 1 are to bid on a single item or to pass its turn. Given the fact that player 2 is better off by obtaining both items in the worst-case scenario than by just acquiring

an item at the current bid price, player 2 will outbid player 1 if it decides to bid on a single item. Thus, the three remaining moves for player 1 are equivalent as they all lead to a utility of 0. Hence, it is highly possible that player 1 passes its turn if player 2 bids on both items. Believing that, player 2 might consider bidding on both items to obtain a utility of 8 which is very risky. Thus, by introducing penalty  $no\_object(x)$  in our selection strategy, we favour the selection of actions where player 1 bids on a single item and does not pass its turn. Hence, this decreases the average reward of the action where player 2 bids on both items and ensures that it stays negative. Ultimately, this discourages player 2 from bidding. Hence, player 2 passes its turn and avoids obvious exposure.

- The fourth term  $risky\_move(x)$  is a penalty term which has been introduced to deter players from bidding on sets of goods which might lead to exposure. These are sets that contain a subset of goods which, if acquired by the player  $i$  bidding at parent node  $y$ , will result in a negative utility. In other words, if player  $i$  decides to pass all its turns until the auction ends after just having bid on a set of items which might lead to exposure, then there is a possibility that  $i$  ends up with a negative utility. The actions leading to such sets are penalised by  $\lambda_i v_i(\{1, \dots, m\})$  with  $\lambda_i \in [0, 1]$  a risk-aversion hyperparameter. Depending on whether  $i$  is the root player  $r$  or not,  $\lambda_i$  either takes the value  $\lambda^r$  or  $\lambda^o$ . The term  $v_i(\{1, \dots, m\})$  acts as a scaling factor. More formally, if  $P^x$  is the price vector at child node  $x$ ,  $r$  is the root player,  $i$  is the player bidding at parent node  $y$  and  $X_x^i$  the set of goods temporarily won by player  $i$  at node  $x$ , then  $X_x^i$  is said to lead to exposure if  $\exists Y \subseteq X_x^i, \sigma_i(Y, P^x) < 0$  and the penalty term is defined as

$$risky\_move(x) = \begin{cases} \lambda^r v_i(\{1, \dots, m\}) & \text{if } X_x^i \text{ can lead to exposure for } i = r \text{ at price } P^x \\ \lambda^o v_i(\{1, \dots, m\}) & \text{if } X_x^i \text{ can lead to exposure for } i \neq r \text{ at price } P^x \\ 0 & \text{otherwise} \end{cases} \quad (4.14)$$

The following example illustrates how the penalty  $risky\_move(x)$  reduces the risk of exposure of the root player. Suppose that  $P = (0, 0)$  and it is player 2's turn to bid in *Example 1* (see Section 2.2.3). Player 2 has the choice between passing its turn, bidding on item 1, bidding on item 2 or bidding on both items. These moves lead respectively to node  $x_0$ ,  $x_1$ ,  $x_2$  and  $x_{1,2}$  in the search tree. As  $\sigma_2(\{1\}, P^{x_1}) = \sigma_2(\{2\}, P^{x_2}) = \sigma_2(\{1\}, P^{x_{1,2}}) = -1 < 0$ , then each bidding move is penalised by  $\lambda^r v_2(\{1, 2\})$ . This deters player 2 from bidding and reduces its risk of exposure.

This selection strategy is adapted to any scale and underlying distribution of rewards.

## 4.4.2 Expansion

The expansion phase consists in choosing which children of the leaf node obtained in the selection phase are expanded to the search tree. In our MCTS, a node is chosen randomly amongst the non-expanded children of the selected leaf node and added to the search tree. The different statistics of this newly added node  $x$  are set the following way:

$$* r_x = 0$$

$$* n_x = 0$$

$$* a_x = +\infty$$

$$* c_x = -\infty$$

## 4.4.3 Rollout strategy

In the rollout phase, moves are played starting from the newly added node using a rollout strategy until the game ends in order to simulate the outcome of the game from this particular node. As the game considered is a  $n$ -player game, the outcome is a vector of size  $n$  where each index corresponds to the utility obtained by each player at the end of the simulation. To simulate a game, the default rollout strategy used in MCTS is usually to play randomly. However, in the considered game, playing randomly will lead to poor quality sampling. Indeed, unlike in Chess or Go, each action can drastically change the outcome of the game. For example, a player who does not wish to buy any items can bid on all of them by playing randomly and end up with an absurd negative payoff. Moreover, the probability that everybody passes their turn is inferior to  $\frac{1}{2^{(n-1)m}}$  which leads to extreme high prices and exposure of all players in expectation. Therefore, to guarantee good sampling, we propose a new rollout strategy which is based on our prediction of final prices  $P^*$ .

At the beginning of each rollout phase, we set  $P^{init} = P^* + \eta$  where  $\eta$  is a random variable which follows a bounded uniform distribution  $U([- \varepsilon, \varepsilon]^m)$ . During the simulation, each bidder plays PP with initial prediction  $P^{init}$ . Ties are still broken in favour of smaller subsets but no longer by selecting the goods with the lowest indices. Instead, they are chosen randomly.

The simulation ends when a terminal state of the d-SAA game is reached. By applying noise to the initial prediction and by breaking ties randomly, players bidding behaviours are diversified at each new simulation which improves the quality of sampling of  $MS^\lambda$ .

#### 4.4.4 Backpropagation

The backpropagation phase consists in propagating backwards the outcome of the game obtained in the rollout phase from the newly added node to the root node to update the statistics stored in each selected node. Let  $V$  be the vector of utility obtained in the rollout phase. Let  $x$  be a selected node which has  $y$  as parent node. Let  $i$  be the player playing at node  $y$ . The statistics at node  $x$  are updated as follows:

$$* r_x \leftarrow r_x + V_i$$

$$* n_x \leftarrow n_x + 1$$

$$* a_x \leftarrow \min(a_x, V_i)$$

$$* c_x \leftarrow \max(c_x, V_i)$$

#### 4.4.5 Final move selection

The four above steps run until a computational criteria (time, memory, iteration constraint) is reached. A final move selection is then performed to choose which action to play. We decide to select the move leading to the child with the highest penalised average. More formally, our final move selection chooses to play the action which leads to child  $x$  with the highest quantity  $q_x$ .

$$q_x = \frac{r_x}{n_x} - no\_object(x) - risky\_move(x) \quad (4.15)$$

#### 4.4.6 Transposition table

In order to share statistics between nodes of the search tree representing the same states, we decide to use a transposition table. The hash function is described in Algorithm 3 and uses information about the instance of the d-SAA game, the state represented at the root node of the search tree  $x_0$  and the state being hashed represented at node  $x$ . As a state

in a d-SAA game is defined by the player  $i$  to bid, the current bid price  $P_j$  and the current winner  $A_j$  of each item  $j$ , using these as only inputs of our hash function regarding the state being hashed is sufficient. Our hash function depends on a hyperparameter  $d_{max}$  which corresponds to a strict upper-bound of the maximal depth of the final search tree. In practise, setting  $d_{max}$  to 10 is more than enough. Let  $P^0$  be the bid price vector at root node  $x_0$  and  $P^x$  be the bid price vector at node  $x$ . As a player can only raise the bid price of an item by the bid increment  $\varepsilon$  each round,  $\frac{P_j^x - P_j^0}{\varepsilon} < d_{max}$  if the depth of node  $x$  is strictly lower than  $d_{max}$ .

---

### Algorithm 3 Hash function for d-SAA

---

**Inputs Game:**  $n$  number of players,  $m$  number of goods,  $\varepsilon$  bid increment  
**Inputs Root:** Bid price  $P^0$  at root node  $x_0$   
**Inputs Node:** Bid price  $P^x$ , current allocations  $A^x$  and  $i$  player bidding at node  $x$   
**Hyperparameter:**  $d_{max}$  a strict upper-bound of the maximal depth of the final search tree  
 $step = d_{max} \times n$   
 $h = i \times step^m$   
**for**  $j = 1, 2, \dots, m$  **do**  
    **if**  $A_j^x > 0$  **then**  $h+ = (d_{max} \times (A_j^x - 1) + \frac{P_j^x - P_j^0}{\varepsilon}) step^{j-1}$   
    **end if**  
**end for**  
**return**  $h$

---

If the depth of every node  $x$  in the search tree is strictly lower than  $d_{max}$ , then our hash function assigns a unique value to every different state represented in the search tree. This value is then remapped to correspond to an index of the transposition table. If this value has never been encountered, a new entry is created in our transposition table. By doing so, given the fact that the maximal depth of the final search tree is strictly lower than  $d_{max}$ , our hash function is a perfect hash function as no type-1 error or type-2 error occurs. In practise, this will always be the case.

#### 4.4.7 Interpretation of $MS^\lambda$

Through the different Monte-Carlo simulations,  $MS^\lambda$  is able to perform a cost-benefit analysis of bidding on each set of goods. More precisely,  $MS^\lambda$  judges if it is worthy to form a collusion with the other bidders to keep prices low and, if so, separates the items fairly according to the value functions of each bidder. This tradeoff between competition and collusion is done by comparing the utility obtained by the different possible collusions between bidders, which correspond to the shallow terminal nodes in the search tree, and the estimated utility obtained by competing through the rollout phase of  $MS^\lambda$ . Moreover,

$MS^\lambda$  also uses the results obtained in the rollout phase to estimate the possible exposure of each bidder. This is especially pronounced when bidders use an underestimation of the *frontier prediction of final prices* as initial prediction of their playout strategy due to the discrete uniform distribution as they bid more aggressively.

$MS^\lambda$  uses two hyperparameters  $\lambda^r$  and  $\lambda^o$  presented in (4.14) having both opposite effects on the algorithm's risk-aversion. Indeed, the higher  $\lambda^r$  is, the less the root player is likely to bid on sets of goods which might lead to exposure whereas the higher  $\lambda^o$  is, the higher the root player's opponents are perceived as risk-averse and the more risky bids seem beneficial.

The penalty *no\_object* is introduced to reduce the selection of the unlikely favourable scenario where the root player's opponents pass their turn and potentially end the auction with only undesired objects while the prices are still low. By doing so, our algorithm  $MS^\lambda$  is more robust regarding exposure and is encouraged to collude with others in order to keep prices low.

## 4.5 Experiments

We now analyse the performance of our MCTS bidding strategy  $MS^\lambda$  in a variety of d-SAA games by comparing it to five other strategies:

- An MCTS which is similar to  $MS^\lambda$  but without the penalties in its selection phase. We name it  $MS^{np}$  for  $MS^\lambda$  with no penalties.
- A *UCB* algorithm [Lattimore, 2020] using the same simulations and selection index as our MCTS but without the penalties.
- Straightforward bidding (SB) [Milgrom, 2000] described in Section 2.4.1.
- The *EPE* algorithm [Wellman, 2008] described in Section 2.4.2.2. It is equivalent to EDPE strategy as the game is with complete information.
- The *SCPD* algorithm [Wellman, 2008] described in Section 2.4.3.1.

It is important to note that, in the upcoming results, none of the players are aware of their opponents' strategy.

The hyperparameters used for  $MS^\lambda$  are  $\lambda^r = 0.1$  and  $\lambda^o = 0.1$ . They are obtained by grid-search. The experiments were run on a server consisting of Intel®Xeon®E5-2699 v4 2.2GHz processors. All algorithms were given a maximum of 30 seconds CPU thinking time.

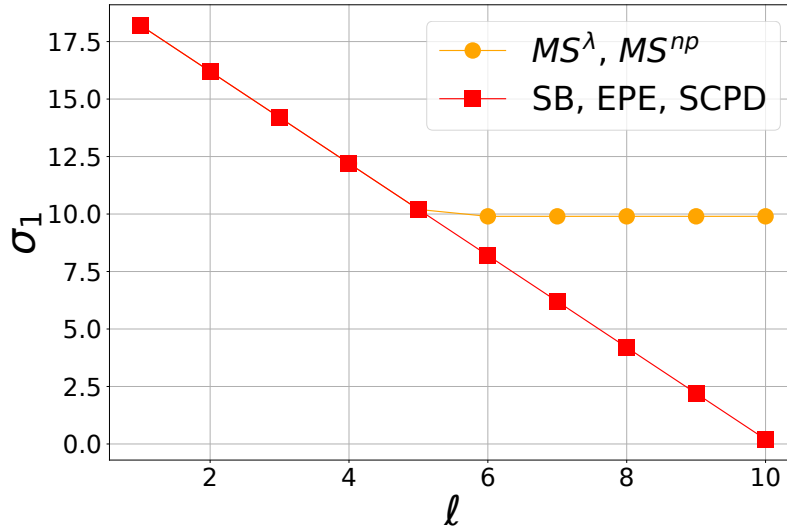
### 4.5.1 Test cases

One of the biggest advantages of  $MS^\lambda$  compared to other existing methods is that it is able to judge pertinently whether it is more beneficial to adopt a demand reduction strategy to keep prices low or to bid straightforwardly on a set of goods. To highlight this capacity, we use an experiment from [Brusco, 2002] in which two players participate in a 2-item auction with additive value functions. The first player values each object as  $h$  and the second values each one as  $0 < \ell \leq h$ . For an infinitesimal bid increment  $\varepsilon$ ,

- If  $h < 2(h - \ell)$ , it is more worthwhile for the first player to bid straightforwardly on both goods. It will then obtain a utility of  $2(h - \ell)$ .
- If  $h > 2(h - \ell)$ , the first player is better off by forming a collusion and conceding an item to its opponent. It will then obtain a utility of  $h$ .

The four algorithms UCB, EPE, SCPD and SB always suggest to bid straightforwardly even if  $h > 2(h - \ell)$  and, thus, never propose demand reduction even in situations where it is highly beneficial. However,  $MS^\lambda$  always adopts the appropriate strategy. We plot in Figure 4.3 the payoff  $\sigma_1$  obtained by player 1 for each strategy given that player 2 plays optimally, i.e., continues to bid on the cheapest item while its bid price is inferior to  $\ell$  and player 1 has not conceded an item to player 2. This example shows that  $MS^\lambda$  chooses the most beneficial strategy, at least in simple environments.

Moreover,  $MS^\lambda$  is capable of avoiding obvious exposure. For instance, in *Example 1* (see Section 2.2.3),  $MS^\lambda$  suggests to player 2 not to bid. EPE and UCB also prevent player 2 from bidding. However, this is not the case for SCPD and SB which expose player 2 by inciting player 2 to bid aggressively.



**Fig. 4.3.:** Evolution of player 1's utility  $\sigma_1$  depending on strategy versus player 2's valuation  $\ell$  in Test experiment [Brusco, 2002] ( $h = 10, \varepsilon = 0.1$ ) given that player 2 plays optimally

## 4.5.2 Extensive experiments

We study the performance of each algorithm mainly through our performance indicators. Each experimental result has been run on 1000 different d-SAA instances. We will focus on d-SAA with  $n = 2, m = 7$  and  $\varepsilon = 1$ . With the exception of our work, all experimental results in the literature are obtained for specific settings of SAA, i.e., using value functions with some specific property such as superadditivity [Goeree, 2014; Wellman, 2008; Reeves, 2005]. Moreover, according to [Goeree, 2014], assigning values to each item and set of items separately has only been undertaken in SAA instances with two items [Zheng, 2012; Brusco, 2002; Brusco, 2009]. Most works prefer to use a one-dimensional type assumption for tractability. For instance, in the case of super-additive value functions, synergies are captured either by adding a known constant to the sum of individual item values or by multiplying this sum by a known multiplier greater than 1. Hence, due generally to the high specificity of settings in experimental results for SAA, it is difficult to conclude on the effectiveness of a method in more generic settings. Therefore, we propose a more general approach to generate value functions by making no additional assumption on its form.



**Setting 1.** Let  $\Gamma$  be an instance of d-SAA with  $n \geq 2$  players,  $m \geq 1$  goods, bid increment  $\varepsilon$  and maximum stand-alone value  $V > 0$ . Each player  $i$  has a general value function  $v_i$  such that  $v_i(\emptyset) = 0$  and, for any set of goods  $X$ , we have

$$v_i(X) \sim U([\max_{j \in X} v_i(X \setminus \{j\}), V + \max_{j \in X} v_i(X \setminus \{j\}) + v_i(\{j\})]) \quad (4.16)$$

where  $U$  is the uniform distribution.

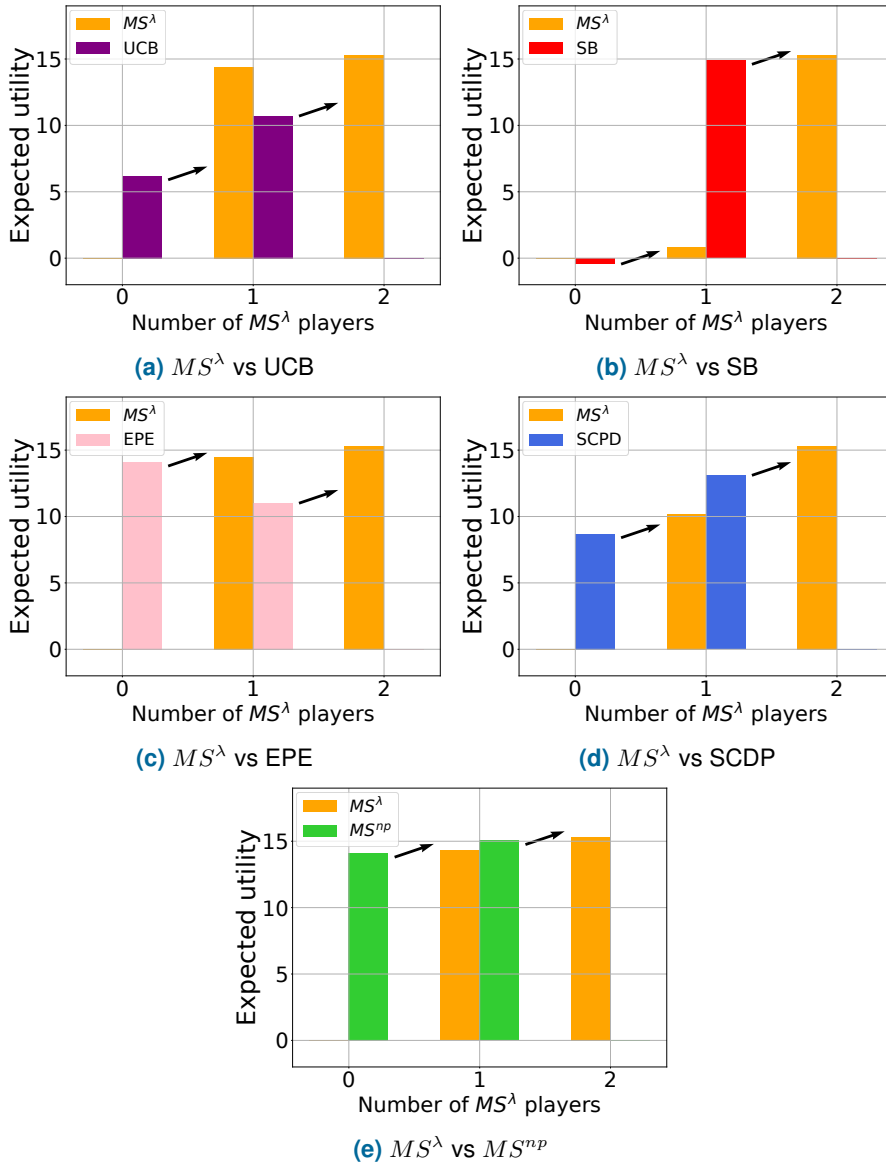
Drawing value functions through a uniform distribution is widely used for creating auction instances [Wellman, 2008; Reeves, 2005]. In our setting, the lower-bound ensures that  $v_i$  respects the *free disposal* [Milgrom, 2000] condition. The upper-bound caps the maximum surplus of complementarity possibly gained by adding an item  $j$  to the set of goods  $X \setminus \{j\}$  by  $V$ . As valuations are always finite, any value function can be represented by our setting for a sufficiently large  $V$ . For  $V = 0$ , only subadditive functions are considered. For  $V > 0$ , goods can either be complements or substitutes. In our experimental results, value functions are generated for each instance as above with  $V = 5$ .

*Remark.* The range of every value function in Setting 1 is contained within  $[0, V(2m - 1)]$ .

As value functions are generated randomly and are associated to a bidder's position, it might be advantageous to play first on average for the 1000 different d-SAA instances. Indeed, as players do not have the opportunity to play at the same bid price, the order in which players submit their bids may have an impact on the auction's outcome. To eliminate such variance and guarantee a fair comparison between two strategies  $A$  and  $B$ , for each d-SAA instance, a game will be run with the first bidder playing  $A$  and the second playing  $B$  and another with the first bidder playing  $B$  and the second playing  $A$ .

#### 4.5.2.1 Expected utility

We first analyse the performance of  $MS^\lambda$  through the same empirical game analysis approach as Wellman et al. [Wellman, 2008] which maps strategy profiles to the average payoff obtained in the 1000 different d-SAA instances for each player. More precisely, we study the symmetric normal form game in expected payoff where each player has the choice between playing our MCTS bidding strategy  $MS^\lambda$  or another specified strategy  $A$ . The resulting empirical games for each possible strategy  $A$  are given in Figure 4.4.



**Fig. 4.4.:** Normal-form payoffs for a d-SAA game with six strategies

It is clear that, in each empirical game, the deviation from UCB, EPE, SCPD, SB or  $MS^{np}$  to  $MS^\lambda$  is always profitable. Thus, the strategy profile  $(MS^\lambda, MS^\lambda)$  is the only Nash equilibrium of the normal-form d-SAA game in expected payoffs with strategy set  $\{MS^\lambda, UCB, EPE, SCPD, SB, MS^{np}\}$ . It is also important to note the significant increase in average payoffs between the strategy profile  $(MS^\lambda, MS^\lambda)$  and the other strategy profiles where all bidders play the same strategy. For instance, in Figure 4.4, the profile  $(MS^\lambda, MS^\lambda)$  has an increase of 108%, 247%, 108% and 175% compared respectively to  $(MS^{np}, MS^{np})$ ,  $(UCB, UCB)$ ,  $(EPE, EPE)$  and  $(SCPD, SCPD)$  in average payoffs. Complete information of

d-SAA enables EPE bidders to share the same expected competitive equilibrium if their tâtonnement process uses the same initial price vector and adjustment parameter. This explains that (EPE,EPE) obtains an expected utility nearly as high as  $(MS^{np},MS^{np})$ . Due to exposure, the profile (SB,SB) obtains a negative expected utility. Hence, both players would have preferred not to participate in the d-SAA initially.

The relative high performance of  $MS^\lambda$  can be explained by three factors: (i) its ability to judge in which situations it is more beneficial to perform demand reduction or to bid competitively as seen in Section 4.5.1; (ii) its ability to perform demand reduction to keep prices low and (iii) its ability to stop bidding on specific sets of goods early on to avoid exposure.

#### 4.5.2.2 Own price effect

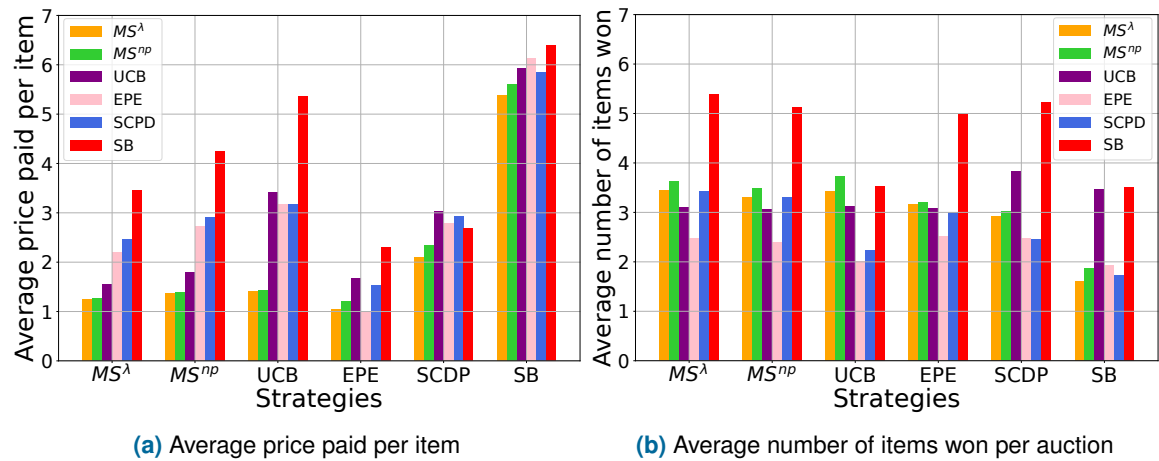


Fig. 4.5.: Own price effect analysis of a d-SAA game with six strategies

Our algorithm  $MS^\lambda$  has the capacity of conceding items to its opponents in order to keep prices low. We highlight this feature in Figure 4.5(a) where we plot the average price paid per item won by each strategy against every strategy displayed on the x-axis.  $MS^\lambda$  obtains the lowest average price paid per item won against every strategy except against EPE. Indeed, each item won is bought at  $\varepsilon$  when both bidders play EPE as they share the same expected competitive equilibrium. This explains the slight underperformance of our algorithm. Nevertheless,  $MS^\lambda$  spends 43.2%, 50%, 55.9%, 24.5% and 12.1% less per item won than EPE against  $MS^\lambda$ ,  $MS^{np}$ , UCB, SCPD and SB respectively. Moreover,  $MS^\lambda$  is competitive and does not just purchase undesired items at relatively low prices. This is highlighted in Figure 4.5(b) where we plot the average number of items won by each strategy against every strategy displayed on the x-axis.  $MS^\lambda$  wins at least 3 items out of

7 against every strategy except SB. In comparison, EPE never wins more than 2.5 items on average against any strategy. Regarding strategy profiles where all bidders play the same strategy, 98.3% of all items are allocated for  $MS^\lambda$  whereas only 71.4% and 70.3% respectively for EPE and SCPD which partially explains their underperformance. The fact that nearly all goods are allocated and are acquired at a small price explains the high performance of the strategy profile  $(MS^\lambda, MS^\lambda)$  compared to the other strategy profiles. Through smart usage of demand reduction, our MCTS bidding strategy  $MS^\lambda$  tackles the own price effect and still remains fairly competitive.

#### 4.5.2.3 Reduction of exposure

As already previously explained, minimising exposure is extremely important in d-SAA. In Figure 4.6(a), we have plotted the exposure frequency of each strategy against every strategy displayed on the x-axis.  $MS^\lambda$  has at most 1.2% of chance of getting exposed against every strategy except SB against which it obtains a similar exposure frequency to SCPD. Moreover, the strategy profile  $(MS^\lambda, MS^\lambda)$  has the remarkable property of never suffering from exposure. It is important to highlight the significant enhancement due to our selection penalties as it is exposed 95.6%, 62.7% and 39.8% less than  $MS^{np}$  against respectively EPE, SCPD and SB. In Figure 4.6(b), we have plotted the expected exposure of each strategy against every strategy displayed on the x-axis.  $MS^\lambda$  expected exposure is at least one order of magnitude below the corresponding SCPD one against  $MS^\lambda$ ,  $MS^{np}$ , UCB and EPE. They are about the same order against SCPD and SB. Moreover,  $MS^\lambda$  generates 78.6%, 66.7%, 61.2% and 37.3% less losses than  $MS^{np}$  against respectively UCB, EPE, SCPD and SB. Thus, in addition to being profitable against all strategies,  $MS^\lambda$  considerably minimises exposure notably through its selection penalties.

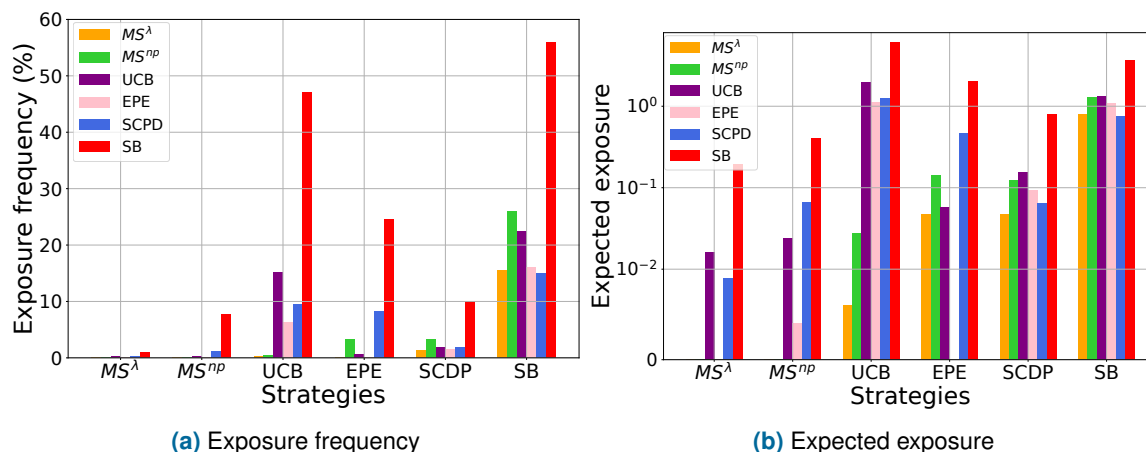


Fig. 4.6.: Exposure analysis of a d-SAA game with six strategies

## 4.6 Conclusion

This chapter introduces the first efficient algorithm that tackles simultaneously the *exposure problem* and the *own price effect* in a simplified version of SAA (d-SAA). Experimental results support the fact that  $MS^\lambda$  largely outperforms state-of-the-art algorithms in d-SAA by obtaining greater expected utility and taking less risks against all other strategies. Its rollout strategy relies on a new concept of solution for the prediction of closing prices named *frontier prediction of final prices*. To the best of our knowledge, it is the first concept that verifies the three following important conditions (1) The concept exists in all auctions (2) It takes into account auctions' particularities (3) The final prediction is independent of a single specific strategy profile. We have shown that such a solution can be easily computed in two-player auctions with two goods when preferences exhibit complementarities through the convergence of a specific sequence. We emit the conjecture that this sequence still converges to one of its solutions for larger auctions. The proof is left for future research.

Contrary to what one might think,  $MS^\lambda$  performs better using a point price prediction rather than a distribution prediction [Wellman, 2008]. This is due to the fact that our MCTS needs to perform a relatively high number of search iterations to be efficient as the action space grows exponentially with the number of goods. However, simulations with distribution price predictors are 10 times more computationally expensive than with point price predictors. Given the relative short CPU thinking time budget, the possibly better sampling quality obtained with distribution predictors doesn't outweigh its relatively poor number of search iterations and, thus, decreases the overall performance of  $MS^\lambda$ .

We believe that the biggest field of improvement for  $MS^\lambda$  is related to its selection phase. Indeed, although the selection penalties have greatly contributed to the reduction of exposure by our algorithm and are adapted to any scale, we think that introducing a more formal statistic of risk-aversion may be more beneficial. However, using quantities such as mean-variance [Liu, 2020] didn't enhance our algorithm's performance. In the next chapter, we propose a new solution that replaces these selection penalties more formally and improves the overall performance of our algorithm.

# Simultaneous Ascending Auction with complete information

## Contents

---

5.1	Simultaneous Ascending Auction model with constraints and complete information (SAA-c) . . . . .	123
5.1.1	Mechanism . . . . .	123
5.1.2	Bidders' modelling . . . . .	123
5.1.3	Representation of the SAA-c game in extensive form . . . . .	124
5.2	Complexities induced by the SAA-c game . . . . .	126
5.2.1	General game properties . . . . .	126
5.2.2	Game complexities . . . . .	126
5.2.3	Strategical complexities . . . . .	128
5.3	Predicting closing prices . . . . .	128
5.3.1	Constrained point-price prediction bidding . . . . .	128
5.3.2	Computing an initial prediction of closing prices . . . . .	129
5.4	SM-MCTS bidding strategy . . . . .	132
5.4.1	Scalarised rewards . . . . .	132
5.4.2	Search tree structure . . . . .	132
5.4.3	Selection . . . . .	134
5.4.4	Expansion . . . . .	135
5.4.5	Rollout . . . . .	136
5.4.6	Backpropagation . . . . .	136
5.4.7	Transposition table . . . . .	136
5.4.8	Final move selection . . . . .	138
5.5	Experiments . . . . .	138
5.5.1	Convergence of sequence $p_t$ . . . . .	139
5.5.2	Test experiments . . . . .	140
5.5.3	Extensive experiments . . . . .	142
5.5.4	Influence of $\alpha$ . . . . .	147
5.6	Conclusion . . . . .	149

---

In the last chapter, we focused on a turn-based deterministic model of SAA with no activity rules or budget constraints. We presented an algorithm  $MS^\lambda$  which obtained very good results on small instances of this model. In this chapter, we consider the original format of SAA presented in Section 1.3.2.1. We extend the work done in the previous chapter to simultaneous moves, random tie-breaking rules, budget constraints, activity rules and larger instances. Nevertheless, we maintain the assumption that the induced bidding game is with complete information. This will be relaxed in the next chapter.

We present the first efficient bidding strategy, named  $SMS^\alpha$ , that tackles simultaneously the four main strategical issues of SAA: the *exposure problem*, the *own price effect*, *budget constraints* and the *eligibility management problem*.  $SMS^\alpha$  is based on Simultaneous Move Monte Carlo Tree Search (SM-MCTS) and uses a hyperparameter  $\alpha$  that allows one to arbitrate between expected utility and risk-aversion. As  $MS^\lambda$ ,  $SMS^\alpha$  relies on a new method based on the convergence of a specific sequence for the prediction of closing prices. This is used to enhance its expansion and rollout phase.

We start by modelling the mechanism of SAA as well as the bidding behaviour of the players by taking into account the new features introduced in this chapter such as simultaneity, random tie-breaking rules, eligibility or budget constraints. We name this model SAA-c. We then present the different complexities induced by the bidding game in SAA-c. We introduce a new method for the prediction of closing prices based on the convergence of a specific sequence and on an extension of strategy PP to constrained environments. We then present our bidding algorithm  $SMS^\alpha$ . Finally, through typical examples taken from the literature and extensive numerical experiments on instances of realistic size, we show that  $SMS^\alpha$  outperforms state-of-the-art algorithms by achieving higher expected utility and better tackling the exposure problem and the own price effect in budget and eligibility constrained environments.

A large majority of the results presented in this chapter comes from the following article *Bidding efficiently in Simultaneous Ascending Auctions with budget and eligibility constraints using Simultaneous Move Monte Carlo Tree Search* [Pacaud, 2023] which is currently reviewed by the journal *IEEE Transactions on Games*.

## 5.1 Simultaneous Ascending Auction model with constraints and complete information (SAA-c)

### 5.1.1 Mechanism

In this chapter, we focus on the original mechanism design of SAA presented in Section 1.3.2.1 where  $m$  indivisible goods are sold via separate and concurrent English auctions between  $n$  players. Bidding occurs in multiple rounds. At the beginning of the auction, the bid price of each item  $j$ , noted  $P_j$ , is set to 0. At each round, players submit their bids simultaneously. Hence, the induced bidding game is with imperfect information. The player having submitted the highest bid on an item  $j$  becomes its temporary winner. If several players have submitted the same highest bid on item  $j$ , then the temporary winner is uniformly chosen at random amongst them. The bid price  $P_j$  of item  $j$  is then set to the highest bid placed on it. The new temporary winners and bid prices are revealed to all players at the end of each round. The auction closes if no new bids have been submitted during a round. The items are then sold at their current bid price to their corresponding temporary winners.

Bidders which do not maintain a certain level of bidding activity are penalised by the activity rule presented in Section 2.1.3, i.e., the number of items temporarily won plus the number of new bids by a bidder can never rise [Goeree, 2014; Milgrom, 2004]. The initial eligibility of each bidder is set to  $m$ .

### 5.1.2 Bidders' modelling

Bidders possess the same characteristics as presented in Section 2.1. Thus, each bidder  $i$  is defined by the three following quantities:

- Value function  $v_i$  (or utility function  $\sigma_i$ )
- Budget  $b_i$
- Eligibility  $e_i$



We assume that these are public knowledge [Szentes, 2003a; Szentes, 2003b; Pacaud, 2022]. Hence, the induced bidding game is with complete information. This assumption will be relaxed in the next chapter.

As it was the case in the last chapter, new bids are constrained to  $P_j + \varepsilon$  where  $\varepsilon$  is a fixed bid increment. This reduction of the bidding space is common in the literature on SAA [Goeree, 2014; Wellman, 2008; Pacaud, 2022]. We make the classical assumption that players won't bid on items that they are currently temporarily winning [Wellman, 2008; Pacaud, 2022]. Hence, in our model, a winner will always pay a price for an item at most  $\varepsilon$  above the highest opponent bid.

We recall that bidders are not allowed to bid on sets of items which exceed either their eligibility or their budget. More formally, if the current bid price vector is  $P$ , a bidder  $i$  temporarily winning a set of items  $Y$  with current eligibility  $e_i$  can bid on a set of items  $X$  if and only if

$$\begin{cases} |X| + |Y| \leq e_i \\ \sum_{j \in X} (P_j + \varepsilon) \leq b_i - \sum_{j \in Y} P_j \end{cases} \quad (5.1)$$

At the end of the auction, the utility obtained by player  $i$  after winning the set of items  $X$  at bid price vector  $P$  is:

$$\sigma_i(X, P) = v_i(X) - \sum_{j \in X} P_j \quad (5.2)$$

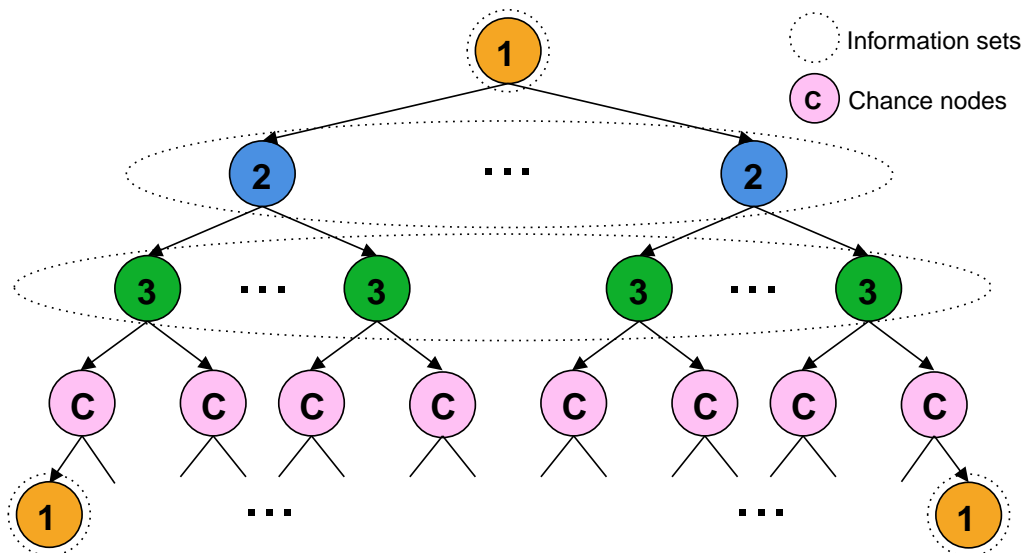
We name this simplified version of SAA with the above bidders' modelling *SAA-c*.

### 5.1.3 Representation of the SAA-c game in extensive form

For multi-round games, the standard representation is the extensive form representation. We distinguish two types of nodes in the extensive form representation of an SAA-c game: *decision nodes* and *chance nodes*. Decision nodes correspond to nodes where the actions are selected by a player and chance nodes where actions are drawn by Nature. More precisely, the decision nodes represent the different states of the game and the chance nodes represent the random draws of temporary winners in case of ties. At each decision node, an outgoing edge represents a feasible bid that can be placed on a set of items satisfying the eligibility and budget constraints by the player bidding at that node. Each decision node or state is defined by five features:

- the identity of the player submitting its bids (we name it the "concerned player")
- the eligibility vector revealed at the end of the last round
- the temporary winner of each item
- the current bid price of each item
- the bids already submitted during the current round

The four first features are common knowledge and the last feature is hidden information for the concerned player. Therefore, all decision nodes which differ only by the last feature belong to the same information set. Thus, unlike the extensive form of a d-SAA game, information sets in the extensive form of a SAA-c game are not all singletons. In Figure 5.1, we represent an SAA-c game between three players with their information sets and chance nodes.



**Fig. 5.1.:** Extensive form of a three player SAA-c game with information sets and chance nodes

## 5.2 Complexities induced by the SAA-c game

### 5.2.1 General game properties

The game induced by SAA-c has the following general game properties:

- n-player game
- Non-zero sum game
- Simultaneous move: Players submit their bids at the same time.
- Stochasticity: This is due to the random tie-breaking rule which selects randomly a temporary winner amongst the players having bid the highest same amount on an item. Hence, a same configuration of bids during a round does not necessarily lead to the same state.
- Imperfect information: This is due to the fact that players bid simultaneously.
- Complete information: All game specificities such as the SAA-c mechanism or the bidders' characteristics are public knowledge.

### 5.2.2 Game complexities

For the sake of simplicity, we compute the information set space complexity and a lower bound of the game tree complexity of an SAA-c game with a given number of rounds  $R$ , unlimited budgets and without any activity rule.

**Theorem 5.1.** *Let  $\Gamma$  be an instance of the SAA-c game with no activity rule. Let  $n$ ,  $m$  and  $R$  be respectively the number of players, the number of items and the number of rounds in  $\Gamma$ . Suppose that all players have unlimited budgets. The number of possible information sets in  $\Gamma$  is:*

$$n(Rn + 1)^m \tag{5.3}$$

*Proof.* Each information set is defined by three components: the player to bid, the temporary winner and bid price of each item. If no player has bid on an item, then it remains unsold and is handed back to the auctioneer. Otherwise, its bid price is included in  $\{\varepsilon, 2\varepsilon, \dots, R\varepsilon\}$  and the item is allocated to one of the  $n$  players. Therefore, the number of different allocations and bid prices of an item in  $\Gamma$  is  $Rn + 1$ . Under the unlimited budget assumption, all items are mutually independent. Thus, the number of different allocations and bid prices for all items is  $(Rn + 1)^m$ . As there are  $n$  different players who can bid, the number of possible information sets is:

$$n(Rn + 1)^m$$

□

As the game is with imperfect information, the above information set space complexity acts as a lower bound of the state space complexity.

**Theorem 5.2.** *Let  $\Gamma$  be an instance of the SAA-c game with no activity rule. Let  $n$ ,  $m$  and  $R$  be respectively the number of players, the number of items, and the number of rounds in  $\Gamma$ . Suppose that all players have unlimited budgets. A lower bound of the game tree complexity of  $\Gamma$  is:*

$$\Omega(2^{m(n-1)R}) \tag{5.4}$$

*Proof.* To prove the above result, we just need to notice that, if we replace the random tie-breaking rule of  $\Gamma$  by a deterministic tie-breaking rule, it reduces its game tree complexity. Indeed, it eliminates chance nodes and reduces the number of paths in its extensive form. Thus, according to the proof of Theorem 4.2, a lower bound of the game tree complexity of  $\Gamma$  with a deterministic tie-breaking rule is  $\Omega(2^{m(n-1)R})$ . Therefore, a lower bound of the game tree complexity of  $\Gamma$  with its original random tie-breaking rule is  $\Omega(2^{m(n-1)R})$ . □

*Example.* An SAA for 12 spectrum licences (5G) between 5 telecommunication companies was held in Italy in 2018 and ended after 171 rounds [European 5G Observatory, 2018]. The number of possible information sets as well as a lower bound of the game tree complexity of the corresponding SAA-c game with no activity rule and unlimited budgets are respectively  $10^{35}$  and  $10^{2470}$ .

Adding budget constraints reduces the information set space complexity and the game tree complexity as some actions can no longer be played and, therefore, some information sets can no longer be reached.

Adding activity rules decreases the game tree complexity as a bidder can no longer bid on a set of items which exceeds its eligibility. However, it increases the information set space complexity as a new feature (eligibility) is added to every information set.

### 5.2.3 Strategical complexities

In addition to the difficulties generated by the high information set space and game tree complexities, the bidding game induced by SAA-c also admits four important strategical issues: *the exposure problem*, *the own price effect*, *budget constraints* and *the eligibility management problem*. All of these issues are fully detailed in Section 2.2.3. In the last chapter, we created an algorithm which only dealt with the two first issues.

## 5.3 Predicting closing prices

$SMS^\alpha$  is based on a SM-MCTS whose expansion and rollout phases rely on the following bidding strategy and prediction of closing prices, i.e., an estimation of the price of each item at the end of the auction.

### 5.3.1 Constrained point-price prediction bidding

We start by extending the definition of strategy PP defined in Section 2.4.2 to budget and eligibility constrained environments.

**Definition 5.1.** In a SAA-c game with  $m$  objects and a current bid price vector  $P$ , a point-price prediction bidder with budget  $b$ , a current eligibility  $e$ , an initial prediction of closing prices  $P^{init}$  and a set of temporarily won items  $Y$  computes the subset of goods

$$X^* = \underset{\substack{X \subset \{1, \dots, m\} \setminus Y \\ \sum_{j \in X \cup Y} \rho_j(P^{init}, P, Y) \leq b \\ |X| + |Y| \leq e}}{\arg \max} \sigma(X \cup Y, \rho(P^{init}, P, Y)) \quad (5.5)$$

breaking ties in favour of smaller subsets and lower-numbered goods. It then bids  $P_j + \varepsilon$  on all items  $j$  belonging to  $X^*$ . The function  $\rho : (P^{init}, P, Y) \rightarrow \mathbb{R}_+^m$  maps an initial prediction of closing prices, a current bid price vector and a set of items temporarily won to an estimation of closing prices. For any item  $j$ , it follows the below update rule:

$$\rho_j(P^{init}, P, Y) = \begin{cases} \max(P_j^{init}, P_j) & \text{if } j \in Y \\ \max(P_j^{init}, P_j + \varepsilon) & \text{otherwise} \end{cases} \quad (5.6)$$

A point-price prediction bidder only considers sets of items *within budget*  $b$  given its prediction of closing prices  $\rho(P^{init}, P, Y)$ , i.e., only sets of items  $X$  such that  $\sum_{j \in X \cup Y} \rho_j(P^{init}, P, Y) \leq b$ . Moreover, it can only bid on sets of items which do not exceed its eligibility  $e$ .

The above extension of PP (which we refer to as PP from now on) to budget and eligibility constrained environments does not change its original properties discussed in Section 2.4.2. For instance, if the initial closing prices are correctly estimated and independent of the bidding strategy, then playing our extension of PP in SAA-c is optimal. Moreover, the accuracy of the initial prediction of closing prices  $P^{init}$  still highly impacts the efficiency of this bidding strategy. For example, it can cause a bidder to bid straightforwardly or to drop out prematurely.

### 5.3.2 Computing an initial prediction of closing prices

As already explained in Section 4.3, several methods exist in the literature for computing an initial prediction of closing prices but they all seem to present some limitations. We especially highlighted three shortcomings. We then presented a simple prediction method based on the convergence of a specific sequence which aims at tackling these three issues. However, the auction considered in the last chapter was deterministic as bidders took turns bidding and did not present any constraints. Thus, we adapt our preceding prediction method to budget constraints, eligibility constraints and stochasticity of SAA-c.

**Conjecture 5.1.** *Let  $\Gamma$  be an instance of an SAA-c game. Let  $f_\Gamma(P)$  be a random variable returning the closing prices of  $\Gamma$  when all bidders play PP with initial prediction  $P$ . The sequence  $p_{t+1} = \frac{1}{t+1} \mathbb{E}[f_\Gamma(p_t)] + (1 - \frac{1}{t+1})p_t$  with  $p_0$  the null vector of prices converges to a unique element  $p^*$ .*

The fact that  $f_\Gamma$  is a random variable comes from the tie-breaking rule which introduces stochasticity in  $\Gamma$ . By taking its expectation  $\mathbb{E}[f_\Gamma(p_t)]$  at each iteration  $t$ , we ensure our deterministic sequence  $p_t$  to always converge to the same fixed point  $p^*$ . Hence, all players using our method share the same prediction of closing prices  $p^*$ . In practice, we perform a Monte-Carlo estimation of  $\mathbb{E}[f_\Gamma(p_t)]$  by simulating many SAA-c games. In small instances, it is possible to obtain a closed-form expression of  $\mathbb{E}[f_\Gamma(p_t)]$  and, from that, prove the convergence of sequence  $p_t$ .

*Example.* Suppose that both players play *PP* with  $P^{init} = p_0$  in *Example 1* (see Section 2.2.3). During the first round, player 1 bids on item 1 and player 2 bids on both items. There is 50% chance that player 1 temporarily wins item 1 and 50% chance that player 2 temporarily wins item 1. If player 1 wins item 1 during the first round, player 2 bids on item 1 during the second round while player 1 passes. In the third round, player 1 bids on item 2 while player 2 passes. In the fourth round, player 2 bids on item 2 while player 1 passes. Hence, the bid price of item 1 (respectively item 2) is odd (respectively even) if temporarily won by player 1. When the bid price  $P = (12, 11)$  and both items are temporarily won by player 2, player 1 drops out of the auction as, by definition of *PP*, it prefers smaller subsets of items for a same predicted utility. If player 2 wins item 1 during the first round, the bid price of item 1 (respectively item 2) is even (respectively even) if temporarily won by player 1. The closing price are then  $P = (11, 11)$ . Therefore,  $f_\Gamma(p_0)$  has 50% chance of returning  $(12, 11)$  and 50% chance of returning  $(11, 11)$ . Hence,  $\mathbb{E}[f_\Gamma(p_0)] = (11.5, 11)$ . By performing a similar analysis, we can show that  $\forall p \in \mathbb{R}^2, \mathbb{E}[f_\Gamma(p)] \in [0, 11.5]^2$  and obtain the following closed-form expression for any  $p \in [0, 11.5]^2$ :

$$\mathbb{E}[f_\Gamma(p)] = \begin{cases} (1, 0) & \text{if } p_1 + p_2 \geq 20 \text{ and } p_1 \leq p_2 \\ (0, 1) & \text{if } p_1 + p_2 \geq 20 \text{ and } p_1 > p_2 \\ (11.5, 11) & \text{if } p_1 + p_2 < 20 \text{ and } p_1 \leq p_2 \\ (11, 11.5) & \text{if } p_1 + p_2 < 20 \text{ and } p_1 > p_2 \end{cases} \quad (5.7)$$

From there, it is easy to show that sequence  $p_t$  converges to  $p^* = (10, 10)$  in *Example 1*. The full proof of the convergence of  $p_t$  to  $p^* = (10, 10)$  with its convergence rates is given in Appendix C.

The general proof of the conjecture is left for future work.

*Remark.* To prove Conjecture 5.1, we think a promising direction is to follow the same direction as in the remark following Conjecture 4.1. However, instead of  $f_\Gamma$ , one should

show that, for all SAA-c instances  $\Gamma$ , the spatial representation of  $\mathbb{E}[f_\Gamma]$  does not admit any cycle.

The following simple example shows one of the main advantages of this new prediction method of final prices compared to the Walrasian price equilibrium or *expected price equilibrium*.

*Example.* Consider an SAA-c game with three players, one item and a bid increment  $\varepsilon = 1$ . The first player values the item at 4.5, the second player values it at 4.5 and the third player values it at 4.6. All three players have a budget of 5.

All Walrasian equilibria can be described with a price vector  $p \in [4.5, 4.6]$  and an allocation vector returning the item to the third player. The tâtonnement processes used to compute the *expected price equilibrium* converges to an element of  $[4.5, 4.6]$  which depends on its adjustment parameter. However, our sequence  $p_t$  converges to the final price of 4. If the first or second player decides to play PP using as initial prediction  $P^{init} \geq 4.5$ , then it drops out immediately of the auction and obtains a utility of 0. However, if it decides to play PP with  $P^{init} = 4$ , then depending on the outcomes of the tie-breaking rule, it either obtains the item for a price smaller than 4 or it does not win the item. Hence, its expected utility is greater than 0. Moreover, in this specific example, playing SB is the optimal strategy. Indeed, if a player is losing temporarily the item and has an eligibility of 1, then it has the choice between either bidding on the item or dropping out definitely of the auction (as its eligibility will be reduced to 0). Hence, if the bid price of the item is less than 4, one should bid if it is temporarily losing the item in order to maximise its expected utility. Moreover, there is no risk of exposure. Thus, playing PP with  $P^{init} < 4.5$  is optimal. By taking into account the auction's mechanism, our prediction method of final prices returns a more accurate initial prediction. Furthermore, as  $f_\Gamma(4) = 4$ , our prediction method also returns the unique self-confirming price prediction of this SAA-c game.

We can discern three main advantages of computing an initial prediction of closing prices through our method than through other methods in the literature. (1) We observe that sequence  $p_t$  converges in all undertaken SAA-c game instances. (2) This method takes into account the auction's mechanism through  $f_\Gamma$ . (3) This prediction of closing price is not based only on the outcomes of a single specific strategy profile. Indeed, depending on the value of  $p_t$ , different strategy profiles are used across iterations. At a fixed iteration  $t$ , a single strategy profile is used to compute  $\mathbb{E}[f_\Gamma(p_t)]$  as the strategy returned by PP only depends on its initial prediction  $P^{init} = p_t$ .



## 5.4 SM-MCTS bidding strategy

In this section, we present our algorithm  $SM S^\alpha$  which is based on Simultaneous Move Monte Carlo Tree Search (SM-MCTS) [Tak, 2014]. It uses a hyperparameter  $\alpha$  which allows a bidder to arbitrate between expected utility and risk-aversion. Before running  $SM S^\alpha$ , we compute our initial prediction of closing prices  $p^*$  as presented in Section 5.3.2. It is used to enhance the expansion and rollout phase of  $SM S^\alpha$ .

### 5.4.1 Scalarised rewards

Maximising the expected utility while minimising the risk of exposure can be antithetical. Indeed, taking risks can either be highly beneficial or lead to exposure depending on how the other players react. To do so, we introduce a new scalarised reward incorporating both targets. Notations used in this part have already been introduced in Section 2.3. For any strategy  $\pi$ , we define:

$$R_\alpha^\pi = (1 + \alpha \mathbb{1}_{R^\pi < 0}) R^\pi \quad (5.8)$$

where  $\alpha$  is a hyperparameter which controls the risk aversion of  $SM S^\alpha$ . Note that

$$\mathbb{E}(R_\alpha^\pi) = \mathbb{E}(R^\pi) + \alpha \mathbf{P}(R^\pi < 0) \mathbb{E}(R^\pi | R^\pi < 0) \quad (5.9)$$

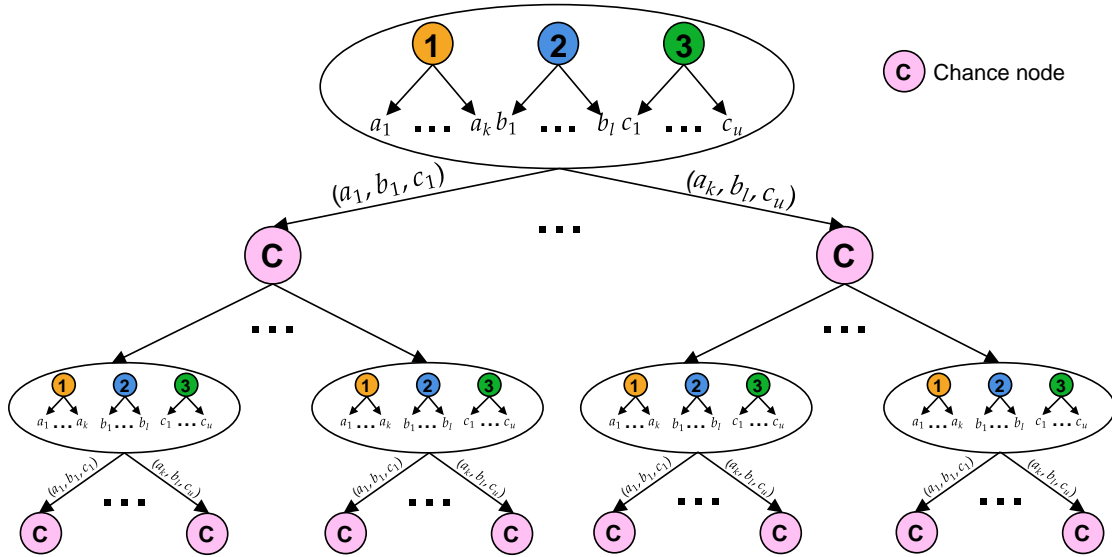
where  $\mathbf{P}(R^\pi < 0) \mathbb{E}(R^\pi | R^\pi < 0)$  is the term corresponding to the losses induced by exposure in Equation 2.6. Moreover, we define for any vector of price  $P$  and any set of items  $X$ ,  $\sigma^\alpha(X, P) = (1 + \alpha \mathbb{1}_{\sigma(X, P) < 0}) \sigma(X, P)$  which is a modified utility taking into account both of our objectives.

The use of a linear scalarization function is a classical approach in multi-objective optimisation, multi-objective reinforcement learning [Barrett, 2008], constrained MDP [Lee, 2018b] or POMDP [Lee, 2018a].

### 5.4.2 Search tree structure

In order to maintain the simultaneous nature of SAA-c in the selection phase of  $SM S^\alpha$ , we use a Simultaneous Move MCTS (SM-MCTS) [Tak, 2014] (Figure 5.2). At each selection step, we select an  $n$ -tuple where each index  $i$  corresponds to the action maximising the selection index of player  $i$  given only its information set. By doing so, bids are selected simultaneously and independently. Each selection step corresponds to a complete bidding

round of SAA-c. Hence, the depth of our search tree corresponds to how many rounds ahead  $SMS^\alpha$  can foresee. The search tree nodes are defined by the eligibility of each bidder, the temporary winner and current bid price of each item. The vertices correspond to players' joint actions. Chance nodes are explicitly included in the search tree to break ties.



**Fig. 5.2.:** SM-MCTS tree structure with explicit chance nodes for a SAA-c game with 3 players

There are three main advantages of using an SM-MCTS instead of an MCTS applied to a serialised game tree, i.e. turning SAA-c into a purely sequential game with perfect information. To simplify our statements, we name *Sequential MCTS* the MCTS applied to a serialised game tree.

The first advantage is that it maintains the simultaneous move nature of SAA-c.

The second advantage is that it does not increase the number of information sets making our learning process more efficient. Thus, less search iterations need to be performed to obtain the same quality of sampling in a SM-MCTS than in a Sequential MCTS as statistics are more often updated. Given the fact that it is generally the rollout phase which is the more time consuming, this results in a higher precision for the statistics of SM-MCTS and, thus, in an increase in performance, especially if the algorithm is only allowed a short period of thinking time.

The third advantage is that the number of selection steps to complete a bidding round of SAA-c is reduced from  $n$  to 1. Thus, the number of players  $n$  is no longer a burden for

planning a bidding strategy over many rounds. However, one may argue that, by reducing the depth of the search tree, we consequently increase its width. Indeed, as the number of joint actions in a bidding round of SAA-c is the same regardless of the search tree structure, it is obvious that performing a bidding round of SAA-c in one step produces a search tree larger than performing a bidding round in  $n$  steps. However, as each index of the  $n$ -tuple is computed independently, this does not result in more basic operations for one selection step of SM-MCTS than for  $n$  steps of Sequential MCTS as shown in the following example.

*Example.* Consider a bidding round of SAA-c between 3 bidders. The first bidder can submit  $K_1$  different bids, the second bidder can submit  $K_2$  different bids and the third bidder can submit  $K_3$  different bids. The number of joint actions is then  $K_1 \times K_2 \times K_3$ . However, for each index of the triplet, an independent selection strategy is run for SM-MCTS. This leads to the same number of basic operations than a Sequential MCTS for 3 selection steps which is roughly  $K_1 + K_2 + K_3$  operations. If the indexes of the joint actions were not selected independently and we had to compute a selection index for each joint action separately, this would have led to roughly  $K_1 \times K_2 \times K_3$  operations and would have indeed been problematic.

### 5.4.3 Selection

At each selection step, players are asked to bid on the set of items which maximises their selection index. The selection phase ends when a terminal state of the SAA-c game or a non-expanded node, i.e. configuration of temporary winners, bid prices and eligibilities not yet added to the search tree, is reached. Our selection index is a direct application of the *Upper Confidence bound applied to Trees* (UCT) [Kocsis, 2006] to scalarised rewards. Unlike usual applications of UCT, the size of the scalarised reward support is unknown so we proceed to an online estimation of it as it was the case for  $MS^\lambda$  in the last chapter. Each player  $i$  chooses to bid on the set of items  $x_i$  with highest score  $q_{x_i}$  at information set  $I_i$ :

$$q_{x_i} = \frac{r_{x_i}^\alpha}{n_{x_i}} + \max(c_{x_i}^\alpha - a_{x_i}^\alpha, \varepsilon) \sqrt{\frac{2 \log(\sum_{x_i'} n_{x_i'})}{n_{x_i}}} \quad (5.10)$$

where  $r_{x_i}^\alpha$  is the sum of scalarised rewards obtained after bidding on  $x_i$  at  $I_i$ ,  $n_{x_i}$  is the number of times player  $i$  has bid on  $x_i$  at  $I_i$ ,  $\varepsilon$  is the bid increment,  $a_{x_i}^\alpha$  is the estimated lower bound and  $c_{x_i}^\alpha$  is the estimated higher bound of the scalarised reward support when

bidding on  $x_i$  at  $I_i$ . Thus,  $\max(c_{x_i}^\alpha - a_{x_i}^\alpha, \varepsilon)$  acts like the size of scalarised reward support when bidding on  $x_i$  at  $I_i$ .

#### 5.4.4 Expansion

Compared to the last chapter where  $MS^\lambda$  was only applied on small instances, our aim is to apply  $SMSS^\alpha$  on instances of realistic size. However, the exponential growth of the game tree's width with the number of items  $m$  induces a high branching factor and, hence, prevents in-depth inspection of promising branches. Thus, it is necessary to reduce the action space at each information set of the search tree [Świechowski, 2023]. To do so, each time a non-expanded node is added to the search tree, we select a maximum number  $N_{act}$  of promising actions per information set. Passing its turn without bidding on any item is always included in the  $N_{act}$  selected actions. This enables  $SMSS^\alpha$  to obtain shallow terminal nodes in its search tree which correspond to collusions between bidders and, thus, reduces the own price effect. The remaining  $N_{act} - 1$  actions correspond to the moves leading to the  $N_{act} - 1$  highest predicted utilities in strategy  $PP$  with initial prediction  $p^*$ . More formally, for each player  $i$  at information set  $I_i$  temporarily winning set of items  $Y_i$  with eligibility  $e_i$ , the action of bidding on set of items  $X_i$  is selected if  $\sigma_i^\alpha(Y_i \cup X_i, \rho(p^*, P, Y_i))$  is one of the  $N_{act} - 1$  highest values with  $P$  the current bid price. Only sets of items  $X_i$  verifying

$$\begin{cases} |X_i| + |Y_i| \leq e_i \\ \sum_{j \in X_i \cup Y_i} \rho_j(p^*, P, Y_i) \leq b_i \end{cases} \quad (5.11)$$

are considered. Statistics for each action are then initialised as follows:

$$* r_{x_i}^\alpha \leftarrow 0$$

$$* n_{x_i} \leftarrow 0$$

$$* a_{x_i}^\alpha \leftarrow +\infty$$

$$* c_{x_i}^\alpha \leftarrow -\infty$$

### 5.4.5 Rollout

From the newly added node, an SAA-c game is simulated until the game ends. Players are asked to bid at each round of the rollout. As already explained in Section 4.4.3, playing randomly in the rollouts generally leads to absurd outcomes. Hence, we propose a similar approach to the one used for the rollout phase of  $MS^\lambda$ .

At the beginning of each rollout phase, we set  $p_i^* = p^* + \eta_i$  with  $\eta_i \sim U([- \varepsilon, \varepsilon]^m)$ . Each player  $i$  then plays  $PP$  with initial prediction of closing prices  $p_i^*$  during the entire rollout. Noise is added to our initial prediction  $p^*$  to diversify players' bidding strategy and, hence, improve the quality of our sampling. At the end of the rollout, an  $n$ -tuple is returned corresponding to the scalarised utility obtained by each player.

### 5.4.6 Backpropagation

The results obtained during the rollout phase are propagated backwards to update the statistics of the selected nodes. Let  $V_i^\alpha$  be the scalarised utility obtained by player  $i$  at the end of the rollout. Let  $x_i$  be the set of items on which player  $i$  bidded at information state  $I_i$  for one of the selected nodes. The statistics stored for  $I_i$  are updated as follows:

$$* r_{x_i}^\alpha \leftarrow r_{x_i}^\alpha + V_i^\alpha$$

$$* n_{x_i} \leftarrow n_{x_i} + 1$$

$$* a_{x_i}^\alpha \leftarrow \min(a_{x_i}^\alpha, V_i^\alpha)$$

$$* c_{x_i}^\alpha \leftarrow \max(c_{x_i}^\alpha, V_i^\alpha)$$

### 5.4.7 Transposition table

Transposition tables are a common search enhancement, already presented in Section 3.9, used to considerably reduce the size of the search tree and improve performance of MCTS within the same computational budget [Childs, 2008]. By using such tables, we prevent the expansion of redundant nodes in our search tree and share the same statistics between transposed information states. This results in a significant improvement in performance of  $SMS^\alpha$  for the same amount of thinking time.

To identify each information set in the search tree, our hash function is based on two functions  $h_1$  and  $h_2$ . The first returns a different integer for each combination of bid prices and allocations. The second returns a different integer for each eligibility vector. Hence, our hash function assigns a unique value to each information set in the search tree. More precisely, due to computational constraints, we can only assign a unique value for every node in the search tree with a depth lower than  $R_{max}$ .  $R_{max}$  is a hyperparameter corresponding to an upper bound of the maximal depth (or rounds) in the final search tree. An example of function  $h_1$  assigning a different integer for each combination of bid prices and allocations in a search tree of maximal depth  $R_{max}$  is given in Algorithm 4. It uses as inputs the bid price vector  $P^0$  at the root of the search tree, the bid price vector  $P$  and the temporary winner  $A_j$  of each item  $j$  at a given node. If  $A_j = 0$ , then item  $j$  is temporarily allocated to the auctioneer.

---

**Algorithm 4** Example of function  $h_1$

---

**Inputs Game:**  $n, m, \varepsilon$   
**Inputs Root Node:** Bid price vector  $P^0$   
**Inputs Node:** Bid price vector  $P$ , Allocation vector  $A$   
**Hyperparameter:**  $R_{max}$   
 $h = 0$   
 $step = R_{max} \times n$   
**for**  $j = 1, 2, \dots, m$  **do**  
    **if**  $A_j > 0$  **then**  $h += (R_{max} \times (A_j - 1) + \frac{P_j - P_j^0}{\varepsilon}) step^{j-1}$   
    **end if**  
**end for**  
**return**  $h$

---

In Section 4.4.6, as players took turns bidding, it was necessary to take into account the next player to bid to identify a node in the search tree. However, as players bid simultaneously in SAA-c and given the SM-MCTS search tree structure of  $SM S^\alpha$ , adding this extra information to determine a search tree node is pointless.

In practice, given the thinking time constraints in our experimental results, choosing  $R_{max} = 10$  is more than sufficient to guarantee a final search tree with maximal depth lower than  $R_{max}$ . Hence, our hash function acts as a perfect hash function as no type-1 error or type-2 error occurs [Zobrist, 1990].

### 5.4.8 Final move selection

The final move which is returned by  $SMS^\alpha$  is the action which maximises the player's expected scalarised reward at the root node. More formally,  $SMS^\alpha$  returns  $\arg \max_{x_i} \frac{r_{x_i}^\alpha}{n_{x_i}}$  for player  $i$ .

## 5.5 Experiments

In this section, we start by analysing the convergence rates of sequence  $p_t$ , notably through *Example 1* (see Section 2.2.3). Then, we show that our algorithm  $SMS^\alpha$  largely outperforms state-of-the-art existing bidding algorithms presented in Section 2.4 in SAA-c, mainly by tackling own price effect and exposure more efficiently. This is first shown through typical examples taken from the literature and, then, through extensive experiments on instances of realistic size. We compare  $SMS^\alpha$  to the following four strategies:

- Our MCTS bidding algorithm  $MS^\lambda$  presented in the last chapter which relies on two risk-aversion hyperparameters  $\lambda^r$  and  $\lambda^o$ .
- The *EPE* algorithm [Wellman, 2008] described in Section 2.4.2.2. It returns the same strategy as EDPE as the game is with complete information.
- The *SCPD* algorithm [Wellman, 2008] described in Section 2.4.3.1.
- Strategy *SB* [Milgrom, 2000] described in Section 2.4.1.

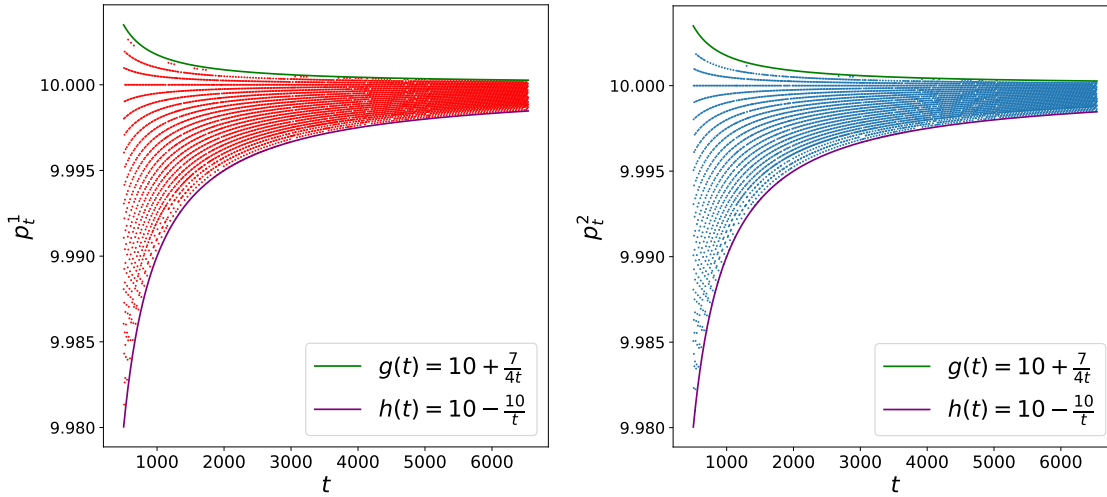
The four strategies  $MS^\lambda$ , EPE, SCPD and SB (equivalent to *PP* with  $P^{init} = 0$ ) initially rely on the definition of perceived-price bidding strategy for unconstrained environments (see Definition 2.1). We extend them to budget and eligibility constrained environments in the same way as it is done in Definition 5.1. In all experiments, none of the bidders are aware of their opponents' strategy.

Each algorithm is given respectively 150 seconds of thinking time. Initial predictions of closing prices are done offline before the auction starts and, therefore, are excluded from the thinking time. This step usually takes a few minutes. All experiments are run on a server consisting of Intel®Xeon®E5-2699 v4 2.2GHz processors. In all upcoming experiments, the hyperparameter  $\alpha$  of  $SMS^\alpha$  takes the value 7 and the risk-aversion hyperparameters

$\lambda^r$  and  $\lambda^o$  of  $MS^\lambda$  both take the value 0.025. These hyperparameters are obtained by grid-search. The maximum number of expanded actions per information set  $N_{act}$  of  $SMS^\alpha$  is set to 20.

### 5.5.1 Convergence of sequence $p_t$

One of the main advantages of using our method to compute an initial prediction is the convergence of sequence  $p_t$ . Even though, this convergence has only been observed and not proven in general cases, it is possible to derive rates of convergence in small instances. For instance, in *Example 1*, it is proved in Appendix C that  $\forall t \geq 1$ ,  $p_t$  belongs to the diamond defined by the points  $(10 - \frac{10}{t}, 10 - \frac{9}{t})$ ,  $(10 - \frac{9}{t}, 10 - \frac{10}{t})$ ,  $(10 + \frac{7}{4t}, 10 + \frac{3}{4t})$  and  $(10 + \frac{3}{4t}, 10 + \frac{7}{4t})$  which converges to  $p^* = (10, 10)$ . We represent in Figure 5.3 both sequences  $p_t^1$  and  $p_t^2$  with their respective lower bound and upper bound.



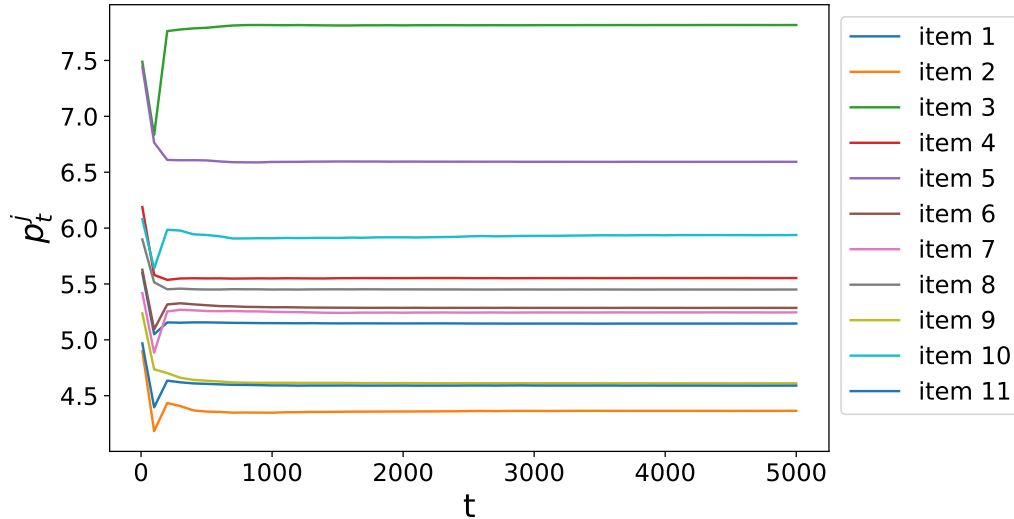
**Fig. 5.3.:** Convergence of sequence  $p_t^1$  and  $p_t^2$  with their respective upper bound  $g(t)$  and lower bound  $h(t)$  in the SAA-c game of *Example 1*.

In larger instances, we observe similar rates of convergence. However, computing such bounds seems unrealistic as obtaining a closed-form expression of  $\mathbb{E}[f_\Gamma(p_t)]$  seems untractable. In Figure 5.4, we plot the value of  $p_t^j$  for every item  $j$  in one of the SAA-c games undertaken in our extensive experiments with  $n = 4$  and  $m = 11$ . We can see that  $p_t^j$  converges after a few hundred iterations for every item  $j$ . It is representative of our observations of sequence  $p_t$  in more complex environments.

*Remark.* As a Monte-Carlo estimator of  $\mathbb{E}[f_\Gamma(p_t)]$  is used for the computation of  $p_{t+1}$  in practice, if sampling is insufficient, it is possible that the sequence  $p_t$  does not always



converge to the same limit  $p^*$  due to the stochasticity of the SAA-c bidding game. Hence, to minimise the risk that  $SM S^\alpha$  players do not share the same limit, it is recommended to sample at least a few thousand games per iteration in complex environments.



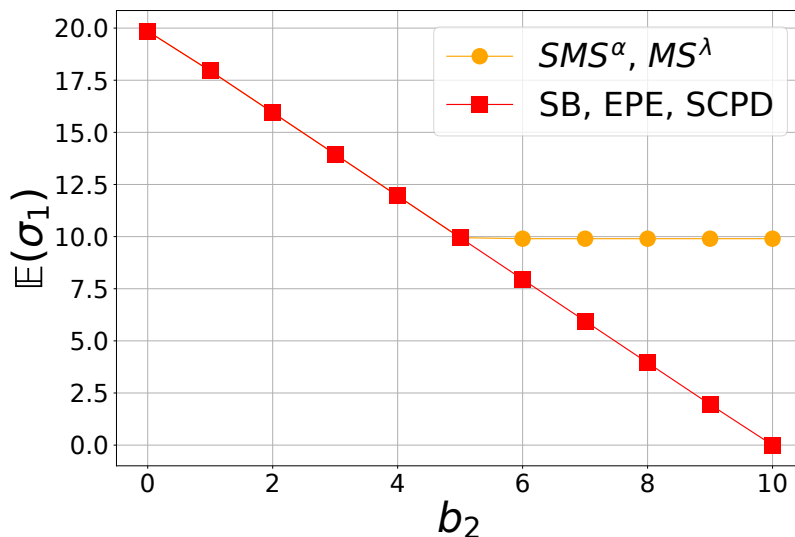
**Fig. 5.4.:** Convergence of sequence  $p_t^j$  for every item  $j$  in an SAA-c game generated from our extensive experiments ( $n = 4$ ,  $m = 11$  and  $\varepsilon = 1$ )

## 5.5.2 Test experiments

One of the greatest advantages that MCTS methods have over other bidding algorithms is the capacity to judge pertinently in which situations adopting a demand reduction strategy is more beneficial. Indeed, through the use of its search tree, an MCTS method is capable of determining if it is more profitable to concede items to its opponents to keep prices low or to bid greedily. To highlight this feature, we propose the following experiment in a 2-item auction between two players with additive value functions. Each player values each item at  $l = 10$ . Player 1 has a budget  $b_1 \geq 20$ . Given that, the optimal strategy for player 2 is to bid on the cheapest item if it is not temporarily winning any item. Otherwise, it should pass. The optimal strategy for player 1 fully depends on its opponent's budget  $b_2$ . For an infinitesimal bid increment  $\varepsilon$ ,

- If  $b_2 \leq \frac{l}{2}$ , player 1's optimal strategy is to play straightforwardly and it obtains an expected utility of  $l - 2b_2$ .
- If  $b_2 \geq \frac{l}{2}$ , player 1's should adopt a demand reduction strategy and it obtains an expected utility of  $l$ .

We plot in Figure 5.5 the expected utility  $\mathbb{E}(\sigma_1)$  of player 1 for each strategy given player 2's budget  $b_2$ . The three algorithms SB, EPE, SCPD always suggest to player 1 to bid greedily and never propose a demand reduction strategy even when it is highly profitable ( $b_2 > \frac{1}{2}$ ). However, both MCTS methods perfectly adopt the appropriate strategy. This experiment highlights the fact that  $SMS^\alpha$  selects the most profitable strategy and tackles own price effect, at least in simple budget and eligibility constrained environments.



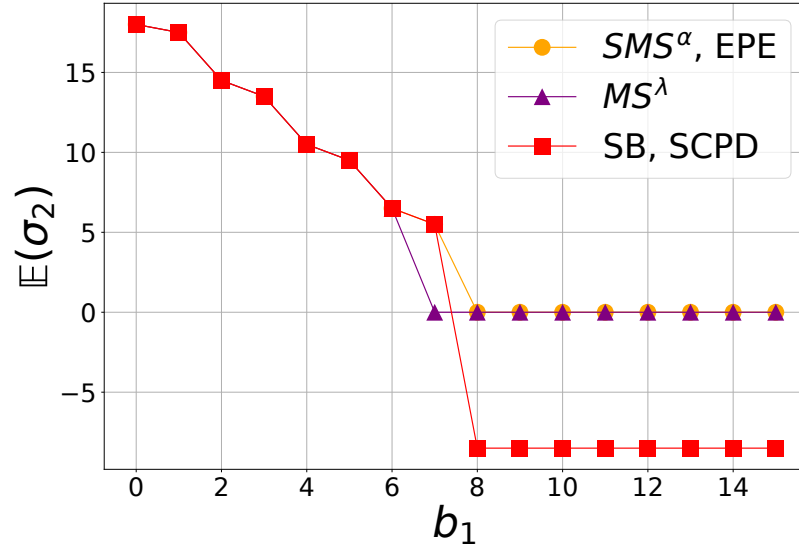
**Fig. 5.5.:** Evolution of player 1's expected utility  $\mathbb{E}(\sigma_1)$  depending on strategy versus player 2's budget  $b_2$  given that player 2 plays optimally ( $\varepsilon = 0.1$ ).

Furthermore,  $SMS^\alpha$  is capable of avoiding obvious exposure. To highlight this feature, we use the SAA-c game presented in *Example 1* (see Section 2.2.3) where player 2's budget  $b_2 = 16$ . The optimal strategy for player 1 is to play straightforwardly. Similarly to the preceding experiment, the optimal strategy for player 2 fully depends on its opponent's budget  $b_1$ .

- If  $b_1 < 8$ , player 2's optimal strategy is to play straightforwardly.
- If  $b_1 \geq 8$ , player 2's optimal strategy is to drop out of the auction to avoid exposure.

We plot in Figure 5.6 the expected utility  $\mathbb{E}(\sigma_2)$  of player 2 for each strategy given player 1's budget  $b_1$ . The two algorithms SCPD and SB always suggest to player 2 to bid straightforwardly leading player 2 to exposure when  $b_1 \geq 8$ .  $MS^\lambda$  never leads player 2 to exposure. However, it suggests to drop out prematurely of the auction in some situations with no risk of exposure due to its selection penalty *risky\_move* and, hence, incurs a loss of easy profit ( $b_1 = 7$ ).  $SMS^\alpha$  and *EPE* perfectly adopt the optimal strategy. This experiment highlights

the fact that  $SMS^\alpha$  perfectly adopts the most profitable strategy and tackles efficiently exposure, at least in simple budget and eligibility constrained environments.



**Fig. 5.6.:** Evolution of player 2's expected utility  $\mathbb{E}(\sigma_2)$  depending on strategy versus player 1's budget  $b_1$  given that player 1 plays optimally in *Example 1* ( $\varepsilon = 1$ ).

### 5.5.3 Extensive experiments

In this section, we study instances of realistic size with  $n = 4$  and  $m = 11$ . Each experimental result has been run on 1000 different SAA-c instances. As explained in Section 4.5.2, all experimental results in the literature are obtained for specific settings and, thus, it is difficult to conclude on the effectiveness of a method in more generic settings. Hence, to overcome this issue, we adopt the same general approach as in the last chapter to generate value functions. Budgets are drawn randomly.

**Setting 2.** Let  $\Gamma$  be an instance of SAA-c with  $n$  bidders,  $m$  items and bid increment  $\varepsilon$ . Each player  $i$  has a budget  $b_i \sim U([b_{min}, b_{max}])$ . Let  $V$  be the maximum surplus of complementarity gained by obtaining an extra item. Its value function  $v_i$  is built as follows:  $v_i(\emptyset) = 0$  and, for any set of goods  $X$ ,

$$v_i(X) \sim U([\max_{j \in X} v_i(X \setminus \{j\}), V + \max_{j \in X} v_i(X \setminus \{j\}) + v_i(\{j\})]) \quad (5.12)$$

with  $U$  the uniform distribution.

In our experimental results, value functions and budgets are generated for each instance as above with  $\varepsilon = 1$ ,  $b_{min} = 10$ ,  $b_{max} = 40$  and  $V = 5$ .

In the upcoming analysis, the average price paid per item won, the ratio of items won, the expected exposure and the exposure frequency are obtained by confronting a strategy  $A$  to a strategy  $B$ . To facilitate our study, each measure of  $A$  against  $B$  is obtained by averaging the results obtained for the three following strategy profiles:  $(A,B,B,B)$ ,  $(A,A,B,B)$  and  $(A,A,A,B)$ . For instance, if  $A = SMS^\alpha$  and  $B = SB$ , the average price paid per item won by  $SMS^\alpha$  in these three strategy profiles is respectively: 5.96, 5.46 and 4.62. Hence, the average price paid per item won by  $SMS^\alpha$  against  $SB$  is 5.35.

### 5.5.3.1 Expected utility

To facilitate our analysis, we study the normal form game in expected utility where each player has the choice between playing  $SMS^\alpha$  or another strategy  $A$ . The same empirical game analysis approach was employed by Wellman et al. in [Wellman, 2008]. More precisely, we map each strategy profile to the estimated expected utility obtained by each player in the 1000 SAA-c instances. The four resulting empirical games for each possible strategy  $A$  are given in Figure 5.7.

For example, in Figure 5.7(b), if all bidders play EPE, each bidder obtains an expected utility of 10.8. In the case of three EPE bidders and one  $SMS^\alpha$  bidder, the  $SMS^\alpha$  bidder obtains an expected utility of 21.5. Hence, if all bidders play EPE, a bidder can double its expected utility by switching to  $SMS^\alpha$ . Therefore, deviating to  $SMS^\alpha$  is profitable if all bidders play EPE. This is also the case for the three other possible deviations in Figure 5.7(b). Hence, in the empirical game where bidders have the choice between playing  $SMS^\alpha$  or EPE, each bidder has interest in playing  $SMS^\alpha$ . We can clearly see that all deviations to  $SMS^\alpha$  are also strictly profitable in the three other empirical games. Hence, in each empirical game, a bidder should play  $SMS^\alpha$  to maximise its expected utility. Therefore, the strategy profile  $(SMS^\alpha, SMS^\alpha, SMS^\alpha, SMS^\alpha)$  is a Nash equilibrium of the normal-form SAA-c game in expected utility with strategy set  $\{SMS^\alpha, MS^\lambda, EPE, SCPD, SB\}$ .

Moreover, the strategy profile where all bidders play  $SMS^\alpha$  has a significantly higher expected utility than any other strategy profile where all bidders play the same strategy. This is mainly due to the fact that  $SMS^\alpha$  tackles efficiently the own price effect. For instance, in Figure 5.7, the expected utility of the strategy profile where all bidders play  $SMS^\alpha$  is

respectively 1.13, 1.68 and 3.94 times higher than the ones where all bidders play EPE,  $MS^\lambda$  and SCPD.

The fact that the expected utility obtained by the strategy profile where all bidders play EPE is relatively close to the one where all bidders play  $SMS^\alpha$  can be explained as follows. To compute their expected price equilibrium as initial prediction of closing prices, all EPE bidders in our experiments share the same initial price vector and adjustment parameter in their tâtonnement process. This tâtonnement process is independent of the auction's mechanism and only relies on the estimated valuations of the players. Hence, as SAA-c is a game with complete information, all EPE bidders share the same initial prediction of closing prices and can therefore split up the items between them more or less efficiently.

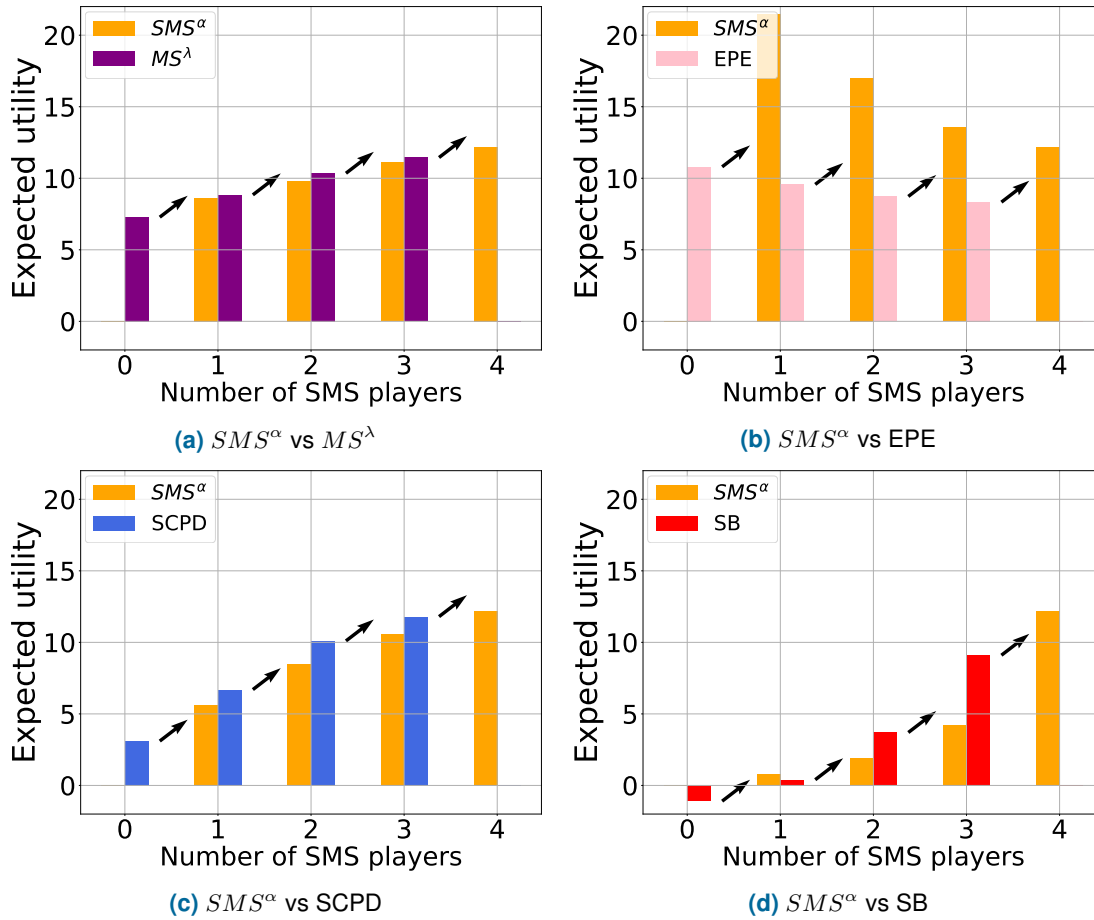
Not all algorithms have the ability of achieving good coordination between bidders. For instance, the strategy profile where all bidders play SB leads to a negative expected utility. Hence, in this specific case, bidders would have preferred not to participate in the auction. This highlights the fact that playing SB is a very risky strategy and mainly leads to exposure.

The high performance of  $SMS^\alpha$  is mostly due to the three following factors:

- its ability to judge if performing demand reduction or bidding greedily is more beneficial given each bidder's budget and eligibility.
- its ability to tackle the own price effect without putting itself in a vulnerable position because of eligibility constraints.
- its ability to avoid exposure in a budget and eligibility constrained environment.

### 5.5.3.2 Own price effect

To analyse own price effect, we plot in Figure 5.8(a) the average price paid per item won by each strategy  $A$  against every strategy  $B$  displayed on the x-axis. For instance, if  $A = SMS^\alpha$  and  $B = EPE$ , the average price paid per item won by  $SMS^\alpha$  against EPE is 1.53. It corresponds to the orange bar above index EPE on the x-axis. If  $A = EPE$  and  $B = SMS^\alpha$ , then the average price paid per item won by EPE against  $SMS^\alpha$  is 2.53. It corresponds to the pink bar above index  $SMS^\alpha$  on the x-axis.



**Fig. 5.7.:** Normal-form expected utility for a SAA-c game with five strategies

In Figure 5.8(a), we can clearly see that  $SMS^\alpha$  acquires items at a lower price in average than the other strategies against  $SMS^\alpha$ , SCPD and SB. For instance,  $SMS^\alpha$  spends 13.3%, 17.1%, 44.9% and 49.8% less per item won against SCPD than  $MS^\lambda$ , EPE, SCPD and SB respectively. Moreover, against  $MS^\lambda$  and EPE, only EPE spends slightly less than  $SMS^\alpha$  per item won.

To ensure that  $SMS^\alpha$  bidders do not obtain low average prices by only purchasing undesired items, we plot in Figure 5.8(b) the ratio of items won by playing each strategy  $A$  against every strategy  $B$  on the x-axis. For instance, if  $A = SMS^\alpha$  and  $B = EPE$ , the ratio of items won by  $SMS^\alpha$  against EPE is 0.31. It corresponds to the orange bar above index EPE on the x-axis in Figure 5.8(b). If  $A = EPE$  and  $B = SMS^\alpha$ , the ratio of items won by EPE against  $SMS^\alpha$  is 0.18. It corresponds to the pink bar above index  $SMS^\alpha$  on the x-axis in Figure 5.8(b). We see that each  $SMS^\alpha$  bidder obtains at least one fifth of the items in average against every strategy except against SB. Hence,  $SMS^\alpha$  is competitive.

Regarding strategy profiles where all bidders play the same strategy, the one corresponding to  $SMS^\alpha$  has an average price paid per item won 1.70, 2.35 and 2.98 times lower than  $MS^\lambda$ , SCPD and SB respectively. Moreover, by looking at Figure 5.8(b), we can see that all items are allocated when all bidders play  $SMS^\alpha$ . Being capable of splitting up all items at a relatively low price explains why the expected utility of the strategy profile where all bidders play  $SMS^\alpha$  is significantly higher than the ones where all bidders play a same other strategy. Only obtaining items at a low price is not sufficient. For instance, when all bidders play EPE, the average price paid per item won is 1.6 times lower than when all bidders play  $SMS^\alpha$ . However, only 72% of all items are allocated. Hence, this strategy profile achieves a lower expected utility than if all bidders had played  $SMS^\alpha$ .

Moreover, the fact that the average price per item won when all bidders play EPE is relatively close to  $\varepsilon$  raises an important strategical issue. Indeed, to obtain such a low price, EPE bidders drastically reduce their eligibility during the first round without considering the fact that they might end up in a vulnerable position. Hence, an EPE bidder can easily be deceived. This explains why a bidder doubles its expected utility if it decides to play  $SMS^\alpha$  instead of EPE when all its opponents are playing EPE in Figure 5.7(b). After the first round,  $SMS^\alpha$  easily takes advantage of the weak position of its opponents. By gradually decreasing its eligibility, a  $SMS^\alpha$  bidder tackles efficiently the own price effect and avoids putting itself in vulnerable positions.

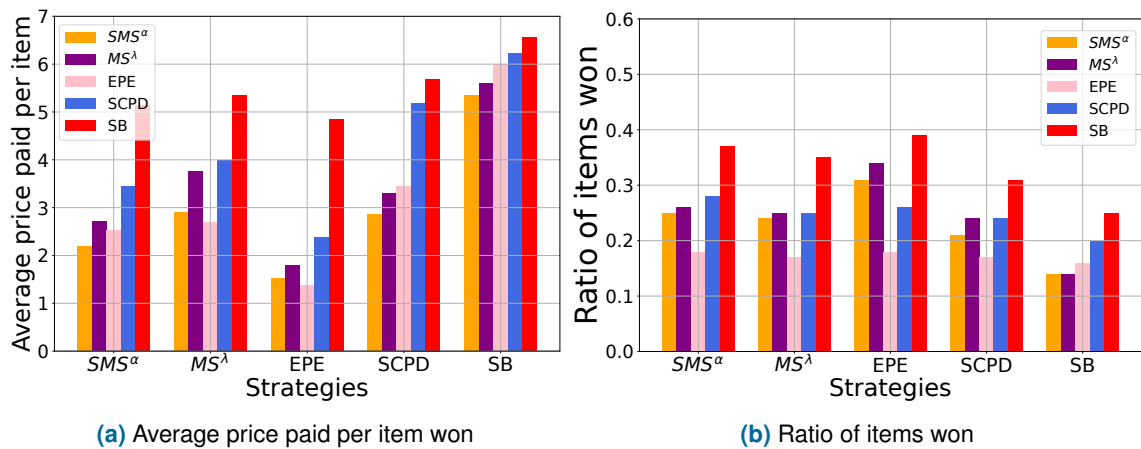


Fig. 5.8.: Own price effect analysis for a SAA-c game with five strategies

### 5.5.3.3 Exposure

To analyse exposure, we plot in Figure 5.9(a) the expected exposure of each strategy  $A$  against every strategy  $B$  displayed on the x-axis. Similarly, we plot in Figure 5.9(b) the exposure frequency of each strategy  $A$  against every strategy  $B$  displayed on the x-axis.

For instance, if  $A = SMS^\alpha$  and  $B = SB$ , the expected exposure and exposure frequency of  $SMS^\alpha$  against  $SB$  are respectively 0.07 and 4.4%. They both correspond respectively to the orange bar above index SB on the x-axis in Figure 5.9(a) and Figure 5.9(b).

Firstly, in the situation where all bidders decide to play the same strategy,  $SMS^\alpha$  has the remarkable property of never leading to exposure. This is not the case for the four other strategies. Secondly,  $SMS^\alpha$  is the only strategy which never suffers from exposure against  $MS^\lambda$  and EPE. Thirdly, even against SCPD and SB,  $SMS^\alpha$  is rarely exposed. It has the lowest expected exposure and exposure frequency. For instance,  $SMS^\alpha$  induces 9.3, 4.5, 34 and 90 times less expected exposure against SCPD than  $MS^\lambda$ , EPE, SCPD and SB respectively. Moreover, regarding exposure frequency, by playing  $SMS^\alpha$  a bidder has 6.6, 4, 27.6 and 58.1 times less chance of ending up exposed against SCPD than  $MS^\lambda$ , EPE, SCPD and SB respectively.

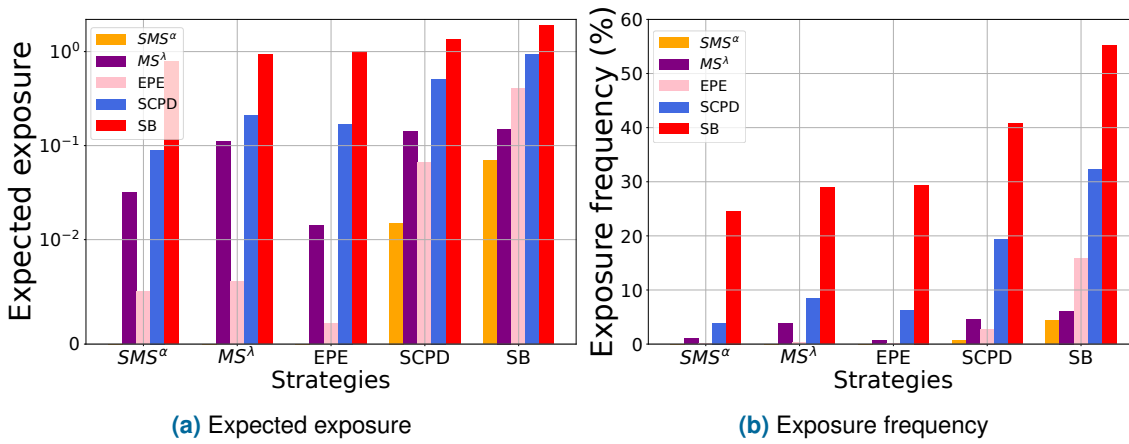


Fig. 5.9.: Exposure analysis for a SAA-c game with five strategies

Hence, not only does  $SMS^\alpha$  achieve higher expected utility than state-of-the-art algorithms but it also takes less risks.

### 5.5.4 Influence of $\alpha$

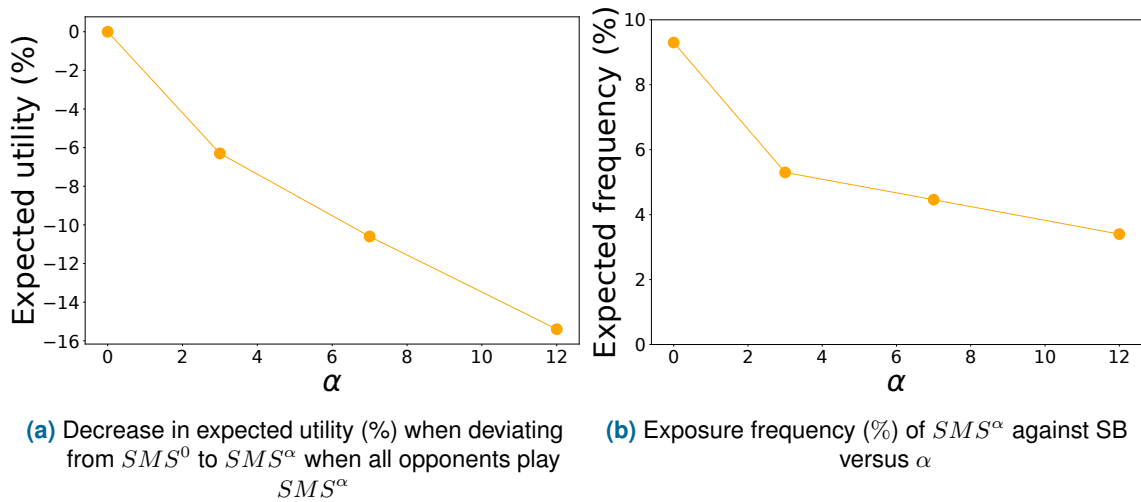
Our strategy  $SMS^\alpha$  is based on a risk-aversion hyperparameter  $\alpha$ . To show its impact on  $SMS^\alpha$ 's performance, we compare  $SMS^\alpha$  for the following values of  $\alpha$ : 0, 3, 7 and 12.

Our first experiment is to study the impact of  $\alpha$  on the expected utility of  $SMS^\alpha$ . We plot in Figure 5.10(a) the relative difference in expected utility between playing  $SMS^0$  and  $SMS^\alpha$  when all other bidders are playing  $SMS^\alpha$ . We observe that switching from  $SMS^0$  to  $SMS^\alpha$



leads to a loss in expected utility for any value  $\alpha > 0$ . Moreover, this loss is an increasing function of  $\alpha$ . Similar results are obtained for the other three deviations in the empirical game where a bidder has the choice between either playing  $SMS^0$  or  $SMS^\alpha$ . The fact that deviating to the risk-neutral strategy  $SMS^0$  is always profitable and that the relative expected loss incurred by switching to  $SMS^\alpha$  increases with  $\alpha$  is far from surprising. Indeed, by increasing  $\alpha$ , a  $SMS^\alpha$  bidder prefers bidding on sets of items which generate less utility but with less chance of leading to exposure.

To highlight the fact that increasing  $\alpha$  leads to less exposure, we plot the exposure frequency of  $SMS^\alpha$  against SB for different values of  $\alpha$  in Figure 5.10(b). We clearly see that the exposure frequency decreases when  $\alpha$  grows. Indeed, an  $SMS^0$  bidder has respectively 1.8, 2.1 and 2.7 more chance of being exposed against SB than  $SMS^3$ ,  $SMS^7$  and  $SMS^{12}$ . It is worth pointing out that the exposure frequency seems to converge to 0 for higher values of  $\alpha$ .



**Fig. 5.10.:** Impact of  $\alpha$  on  $SMS^\alpha$

Through both of these experiments, we highlight the fact that, when playing  $SMS^\alpha$ , increasing  $\alpha$  reduces the risk of exposure but also decreases one's expected utility. Hence, the hyperparameter  $\alpha$  allows the bidder to arbitrate between expected utility and risk-aversion.

In Table 5.1, we highlight another complementary phenomenon tied to risk-aversion when  $\alpha$  increases. Although we have stated above that the expected utility of a bidder playing  $SMS^\alpha$  decreases when  $\alpha$  increases and all other bidders play the same **fixed** strategy, it is no longer the case if all bidders play  $SMS^\alpha$  (opponents no longer play a fixed strategy as  $\alpha$  increases for every player). Indeed, in Table 5.1, we show that, when all bidders play

$SMS^\alpha$  and that  $\alpha$  increases, their expected utility first steadily increases before reaching a ceiling and then decreases before converging to a certain value. This dynamic can easily be explained through two features: the average price per item won and the ratio of allocated items.

During the first part of the dynamic ( $\alpha \in [0, 12]$ ), the average price per item won drastically decreases while the ratio of allocated items remains constant. Hence, the expected utility of the players increases. The fact that increasing  $\alpha$  enables better tackling of the own price effect is a natural effect of risk-aversion where bidders tend to avoid a rise in price.

During the second part of the dynamic ( $\alpha \in [12, 500]$ ), the average price per item won continues to decrease. However, it does not compensate the decrease in the ratio of items won by a player (many items remain unsold at the end of the auction) and, thus, the expected utility decreases up to a certain value. The fact that the proportion of unsold items increases is also tied to risk-aversion where players no longer take the risk of bidding on sets of items which might get them exposed in later rounds.

$\alpha$	0	3	7	12	500
Expected utility	9.25	11.1	12.2	12.6	8.5
Average price per item won	3.11	2.49	2.21	2.01	1.25
Ratio of allocated items	100%	100%	100%	99.7%	53.2%

**Tab. 5.1.:** Evolution of the expected utility, the average price per item won and the ratio of allocated items when all bidders play  $SMS^\alpha$  and  $\alpha$  increases

We can conclude that, by increasing  $\alpha$ ,  $SMS^\alpha$  tackles more efficiently the exposure problem and own price effect. Thus, it minimises the risk of incurring a loss if bidders do not behave as expected. However, the main drawback is that it decreases one's expected utility, notably by decreasing its ratio of items won. The hyperparameter  $\alpha$  thus allows the bidder to arbitrate between expected utility and risk-aversion.

## 5.6 Conclusion

In this chapter, we introduce the first efficient bidding strategy that tackles simultaneously the *exposure problem*, the *own price effect*, *budget constraints* and the *eligibility management problem* in a simplified version of SAA (SAA-c). Our solution  $SMS^\alpha$  largely outperforms state-of-the-art algorithms on instances of realistic size in generic settings.

It is a SM-MCTS whose expansion and rollout phase relies on a new method for the prediction of closing prices. This method is based on a specific sequence that has the advantage of converging in practice in all undertaken SAA-c instances, taking into account the auction's mechanism and not relying solely on the outcomes of a single specific strategy profile. We introduce scalarised rewards in  $SMS^\alpha$  through a hyperparameter  $\alpha$  giving the freedom to bidders to arbitrate between expected utility and risk-aversion. Increasing  $\alpha$  reduces exposure and own price effect but decreases one's expected utility.

In the next chapter, we relax the assumption that the induced bidding game is with complete information. Thus, bidders are no longer aware of the value functions and budgets of their opponents. We propose three different determinization approaches which heavily rely on  $SMS^\alpha$  to address this issue.

# Simultaneous Ascending Auction with incomplete information

## Contents

---

6.1	Simultaneous Ascending Auction with constraints and incomplete information (SAA-inc)	152
6.1.1	Mechanism	153
6.1.2	Bidders' modelling	153
6.1.3	Representation of SAA-inc in extensive form	154
6.2	Complexities induced by the SAA-inc game	156
6.2.1	General game properties	156
6.2.2	Game complexities	157
6.2.3	Strategical complexities	158
6.3	Integrating EXP3 in $SMS^\alpha$	158
6.3.1	Selection	159
6.3.2	Backpropagation	159
6.3.3	Final move selection	160
6.4	Three determinization approaches	161
6.4.1	Determinization through expectation	162
6.4.2	Separate-tree determinization	163
6.4.3	Single-tree determinization	166
6.5	Generating combinations of profiles	172
6.6	Tracking bid exposure	173
6.7	Type generation	174
6.7.1	Value generation	174
6.7.2	Budget generation	178
6.8	Extensive experiments	179
6.8.1	Benchmark	179
6.8.2	Experimental setting	180
6.8.3	Expected utility	182
6.8.4	Own price effect	188

6.8.5	Exposure . . . . .	193
6.8.6	Lessons learned from the experiments . . . . .	196
6.9	Conclusion . . . . .	197

---

In the last chapter, we presented the first efficient bidding strategy  $SM S^\alpha$  that tackles simultaneously the four main strategical issues of SAA on instances of realistic size. However, the induced bidding game was with complete information. In this chapter, we relax this assumption and consider the same bidding game but with incomplete information. Bidders are no longer aware of the value functions or budgets of their opponents. We propose three different determinization approaches highly inspired by  $SM S^\alpha$  in order to compute efficient bidding strategies in the resulting game.

We start by modelling the bidding behaviour of the players as well as its characteristics by introducing the notion of types. The mechanism is the same as in the last chapter. We name this model *SAA-inc*. We then propose a representation of the SAA-inc game in extensive form and present the different complexities induced by the bidding game. We improve the algorithm  $SM S^\alpha$  of the last chapter by replacing the UCT selection phase with the EXP3 algorithm introduced in Section 3.6. We name the resulting algorithm  $SM S_{EXP3}^\alpha$ . We then present three different determinization approaches to adapt  $SM S_{EXP3}^\alpha$  to the incomplete information framework. For two of these determinization approaches, it is necessary to pick multiple types. Given the fact that we can only pick a few because of time constraints, we propose an easy method to sample value functions and budgets which are representative of one's type distribution. Furthermore, we propose a simple inference method to update one's belief about an opponent's budget through bid exposure. We then present a general experimental framework for generating types for a certain level of certainty. Finally, through extensive experiments, we show that our three determinization approaches significantly outperform state-of-the-art methods by notably better tackling the exposure problem and the own price effect in budget and eligibility constrained environments with incomplete information. Moreover, we propose an analysis of the impact of uncertainty on our three determinization approaches.

## 6.1 Simultaneous Ascending Auction with constraints and incomplete information (SAA-inc)

## 6.1.1 Mechanism

In this chapter, we use the same mechanism as in the last chapter presented in Section 5.1.1.

## 6.1.2 Bidders' modelling

Bidders possess the same characteristics as in the last chapter, i.e., each bidder  $i$  is defined by the three following quantities:

- Value function  $v_i$  (or utility function  $\sigma_i$ )
- Budget  $b_i$
- Eligibility  $e_i$

However, the two first characteristics are no longer public knowledge. Hence, the induced bidding game is with incomplete information. Incomplete information games are often referred to as Bayesian Games [[Harsanyi, 1968](#)].

In this chapter, we decide to use a similar Bayesian setting than in [[Nedelec, 2022](#)] to model the incertitude on one's value function and budget. Thus, players no longer have a point-wise estimate of the value function or budget of their opponents, as it was the case for complete information, but a probability distribution. A widely made assumption is that these distributions are common knowledge between bidders [[Nedelec, 2022](#)]. Thus, we denote by  $\mathcal{F}_i$  and by  $\mathcal{B}_i$ , the value distribution and budget distribution of bidder  $i$ . We design by  $\mathcal{F}_{-i}$  (respectively  $\mathcal{B}_{-i}$ ) the product distribution of all value distributions (respectively budget distributions) of all bidders except  $i$ . As budgets and value functions are assumed to be independent in this thesis, we assume that both distributions are independent. We define in this chapter the type distribution  $\mathcal{T}_i$  of player  $i$  as the joint probability that player  $i$  has a value function of  $v$  and a budget of  $b$ , i.e.,

$$\mathcal{T}_i(v, b) = \mathcal{F}_i(v)\mathcal{B}_i(b) \tag{6.1}$$

We define  $\mathcal{T}_{-i}$  as the joint distribution of all type distributions of all players except  $i$ .

We differentiate three types of Bayesian games depending on the information available to bidders. Those are known as *ex-ante*, *interim* and *ex-post* properties [Nedelec, 2022].

- **Ex-ante properties:** Bidders do not know their type and their opponents' types. Hence, each bidder  $i$  only knows  $\mathcal{T}_i$  and  $\mathcal{T}_{-i}$ .
- **Interim properties:** Each bidder  $i$  knows its own type but does not know its opponents' types, i.e., a bidder  $i$  only knows  $v_i, b_i$  and  $\mathcal{T}_{-i}$ .
- **Ex-post properties:** Bidders know their own type and their opponents' types. This implies that the induced bidding game is with complete information.

In this chapter, we consider the Bayesian game with interim properties.

As it was the case in the two last chapters, we make the two following common assumptions which reduce the bidding space:

- Bids on an item  $j$  are constrained to  $P_j + \varepsilon$  where  $\varepsilon$  is a fixed bid increment.
- Players won't bid on items that they are currently temporarily winning. Hence, in our model, a winner will always pay a price for an item at most  $\varepsilon$  above the highest opponent bid.

We name this simplified version of SAA with the above bidders' modelling *SAA-inc*.

### 6.1.3 Representation of SAA-inc in extensive form

According to [Harsanyi, 1967], a Bayesian game can be seen as a game with complete but imperfect information. Indeed, an equivalent representation of a Bayesian game is a game where Nature first draws each player's type in accordance with their type distribution and, then, the players play in the resulting sub-game with complete information. This representation is called the Bayes-equivalent of the original game. Given that our induced bidding game in SAA-inc is with interim properties, each player after Nature's draw receives partial information concerning its exact budget and value function. However, it does not know the budget or value function of its opponents but just their type distribution  $\mathcal{T}_i$ . Thus, a state of the corresponding SAA-inc game is defined by seven features:

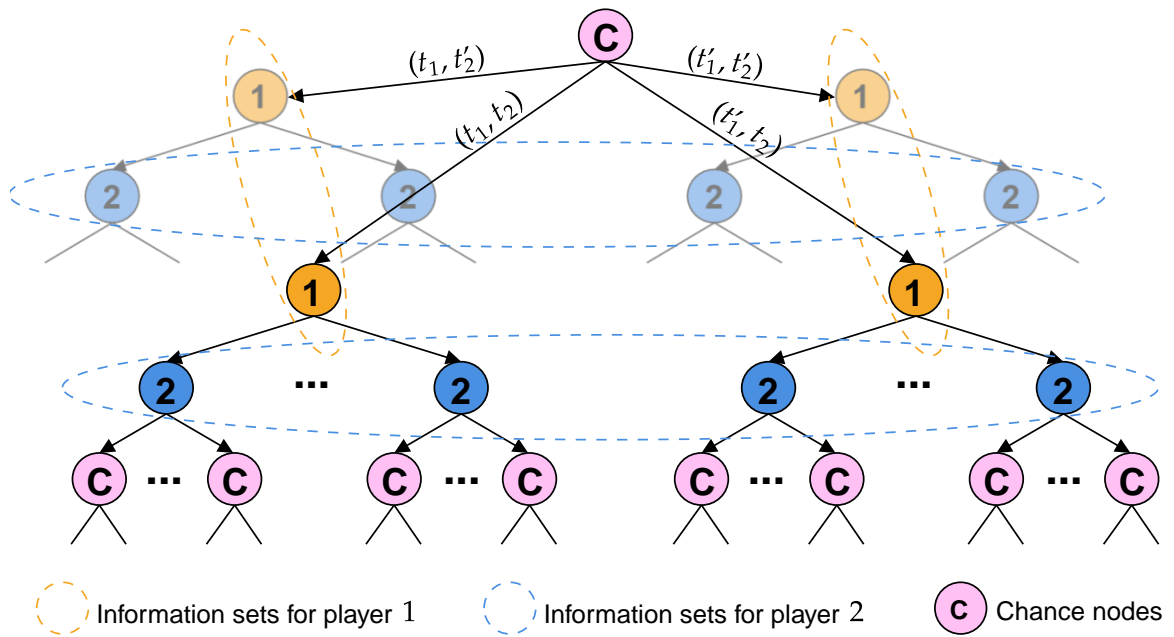
- the identity of the player submitting its bids (we name it the "concerned player")
- the eligibility vector revealed at the end of the last round
- the temporary winner of each item
- the current bid price of each item
- the concerned player's type
- the bids already submitted during the current round
- the types of the concerned player's opponents

The five first features are common knowledge and the two last features are hidden information for the concerned player. Thus, all states which differ only by the last two features belong to the same information set.

In order to theoretically represent the Bayes-equivalent of the SAA-inc game in extensive form, it is necessary that the type distributions are not continuous but discrete. In practice, this will always be the case as the value functions and budgets correspond to monetised values. From now on, we implicitly identify the SAA-inc game with its Bayes-equivalent and use both interchangeably.

An example of SAA-inc game between two players is represented in Figure 6.1. At the beginning of the game, Nature draws a type  $t_i$  for each player  $i$ . Each player knows its type but not its opponent's type. Hence, each player does not know exactly in which sub-game it is playing and, thus, different states across sub-games can belong to a same information state. For instance, all states corresponding to the first round of the auction which share the same concerned player and the same concerned player's type belong to the same information set. Thus, in Figure 6.1, the two nodes corresponding to player 1 with types  $(t_1, t_2)$  and  $(t_1, t'_2)$  belong to the same information set.





**Fig. 6.1.:** Extensive form of a two player SAA-inc game with information sets and chance nodes. The first chance node corresponds to the draw by Nature of each player's type. For example, the couple  $(t_1, t'_2)$  means the first player is type  $t_1$  and the second player is type  $t'_2$ .

## 6.2 Complexities induced by the SAA-inc game

### 6.2.1 General game properties

The game induced by SAA-inc has the following general game properties:

- n-player game
- Non-zero sum game
- Simultaneous move
- Stochasticity
- Imperfect information
- Incomplete information

## 6.2.2 Game complexities

For the sake of simplicity, we compute the information set space complexity and a lower bound of the game tree complexity of an SAA-inc game with a given number of rounds  $R$ , unlimited budgets and without any activity rule.

**Theorem 6.1.** *Let  $\Gamma$  be an instance of the SAA-inc game with no activity rule. Let  $n$ ,  $m$  and  $R$  be respectively the number of players, the number of items and the number of rounds in  $\Gamma$ . Suppose that all players have unlimited budgets. The number of possible information sets in  $\Gamma$  is:*

$$\sum_{i=1}^n |\text{supp}(\mathcal{T}_i)|(Rn+1)^m \quad (6.2)$$

where  $|\text{supp}(\mathcal{T}_i)|$  is the cardinal of the support of the discrete type distribution of player  $i$ .

*Proof.* Each information set is defined by four components: the concerned player, its type (which only corresponds to its value function as all bidders have unlimited budgets), the temporary winner and bid price of each item. Each bidder  $i$  has  $|\text{supp}(\mathcal{T}_i)| = |\text{supp}(\mathcal{F}_i)|$  different types. It has been shown in proof of Theorem 5.1 that the number of different allocations and bid prices for all items in  $\Gamma$  is  $(Rn+1)^m$ . Hence, there are  $|\text{supp}(\mathcal{T}_i)|(Rn+1)^m$  different information sets corresponding to bidder  $i$ . Thus, the total number of possible information sets in  $\Gamma$  is:

$$\sum_{i=1}^n |\text{supp}(\mathcal{T}_i)|(Rn+1)^m$$

□

*Remark.* In the specific case that an instance  $\Gamma$  of SAA-inc game is with complete information, then  $\forall i \in \{1, \dots, n\}, |\text{supp}(\mathcal{T}_i)| = 1$ . Thus, the number of possible information sets in  $\Gamma$  is  $n(Rn+1)^m$  which corresponds to Theorem 5.1.

**Theorem 6.2.** *Let  $\Gamma$  be an instance of the SAA-inc game with no activity rule. Let  $n$ ,  $m$  and  $R$  be respectively the number of players, the number of items, and the number of rounds in  $\Gamma$ . Suppose that all players have unlimited budgets. A lower bound of the game tree complexity of  $\Gamma$  is:*

$$\Omega(2^{m(n-1)R} \prod_{i=1}^n |\text{supp}(\mathcal{T}_i)|) \quad (6.3)$$

where  $|\text{supp}(\mathcal{T}_i)|$  is the cardinal of the support of the discrete type distribution of player  $i$ .

*Proof.* We notice that there are  $\prod_{i=1}^n |supp(\mathcal{T}_i)|$  different type profiles in  $\Gamma$ . Hence, there are  $\prod_{i=1}^n |supp(\mathcal{T}_i)|$  different sub-games which can initially be drawn by Nature. In each of these sub-games, we can lower-bound the game tree complexity by  $\Omega(2^{m(n-1)R})$  as done in the proof of Theorem 5.2. Hence, a lower bound of the game tree complexity of  $\Gamma$  is:

$$\Omega(2^{m(n-1)R} \prod_{i=1}^n |supp(\mathcal{T}_i)|)$$

□

### 6.2.3 Strategical complexities

In addition to the difficulties generated by uncertainty as well as the high information set space and game tree complexities, the bidding game induced by SAA-inc also presents the four main strategical issues of SAA: *the exposure problem*, *the own price effect*, *budget constraints* and *the eligibility management problem*. All of these issues are fully detailed in Section 2.2.3. In the last chapter, we created an efficient algorithm which dealt with these four issues in the specific case of complete information.

## 6.3 Integrating EXP3 in $SMS^\alpha$

Although  $SMS^\alpha$  obtains very good results, we have noticed that these could be improved by replacing the UCT selection phase of  $SMS^\alpha$  with the EXP3 algorithm introduced in Section 3.6. This choice of EXP3 is motivated by the fact that the optimal policy at a simultaneous move node is often mixed [Cowling, 2012a]. Thus, using a selection strategy which returns a mixed policy rather than one which returns a deterministic policy can sometimes lead to a significant improvement of the overall algorithm's performance.

We present hereafter the modifications in the selection phase, the backpropagation phase and the final move selection of  $SMS^\alpha$  to integrate the EXP3 algorithm. The expansion phase and the rollout phase remain unchanged. We name the resulting algorithm  $SMS_{EXP3}^\alpha$ .

### 6.3.1 Selection

During the selection phase, each player  $i$  chooses to bid on the set of items  $x_i$  with probability  $\mathbb{P}_i(x_i)$  at information set  $I_i$ :

$$\mathbb{P}_i(x_i) = \frac{\gamma_{I_i}}{K_{I_i}} + \frac{(1 - \gamma_{I_i})}{\sum_{x'_i} e^{\eta_{I_i}(s_{x'_i}^\alpha - s_{x_i}^\alpha)}} \quad (6.4)$$

where  $s_{x_i}^\alpha$  is an estimate of the sum of scalarised rewards obtained after bidding on  $x_i$  at  $I_i$  over all search iterations where  $I_i$  is encountered in the selection phase,  $K_{I_i}$  is the number of possible different actions which can be selected at  $I_i$  (it is upper-bounded by  $N_{act}$ , the maximum number of promising actions expanded per information set),  $\gamma_{I_i}$  corresponds to the probability of exploring (choosing an action randomly) and  $\eta_{I_i}$  is a hyperparameter of the Gibbs distribution. This equation is slightly different from the one presented in Section 3.6 but is equivalent and more numerically stable [Cowling, 2012a]. We decide to use the same hyperparameters suggested by Auer et al. in [Auer, 2002b],

$$\begin{cases} \gamma_{I_i} = \min(1, \sqrt{\frac{K_{I_i} \ln(K_{I_i})}{(e-1) \sum_{x'_i} n_{x'_i}}} \\ \eta_{I_i} = \frac{\gamma_{I_i}}{K_{I_i}} \end{cases} \quad (6.5)$$

where  $n_{x_i}$  is the number of times player  $i$  bidded on  $x_i$  at  $I_i$  and  $e$  is the base of the natural logarithm.

### 6.3.2 Backpropagation

During the backpropagation phase, we update the statistics of the different selected nodes. Let  $V_i^\alpha$  be the scalarised utility obtained by player  $i$  at the end of the rollout. Let  $x_i$  be the set of items on which player  $i$  bidded at information state  $I_i$  for one of the selected nodes. The statistics stored for  $I_i$  are updated as follows:

$$* s_{x_i}^\alpha \leftarrow s_{x_i}^\alpha + \frac{V_i^\alpha}{\mathbb{P}_i(x_i)}$$

$$* n_{x_i} \leftarrow n_{x_i} + 1$$

By dividing  $V_i^\alpha$  by  $\mathbf{P}_i(x_i)$ ,  $s_{x_i}^\alpha$  acts like an estimate of the sum of scalarised rewards after bidding on  $x_i$  at  $I_i$  for all search iterations where  $I_i$  is encountered in the selection phase and not just over the search iterations where  $x_i$  is selected.

### 6.3.3 Final move selection

When all search iterations have been performed, the algorithm returns a mixed strategy based on the frequencies of visit counts of the expanded moves at the root's information set for the concerned player  $i$ . According to [Teytaud, 2011], before computing these frequencies, withdrawing visits caused by exploration can enhance the overall performance of the algorithm. The average number of visits caused by exploration at information set  $I_i$  can be approximated by

$$\begin{aligned} \sum_{n=1}^{N'_{I_i}} \frac{\gamma_{I_i}}{K_{I_i}} &\approx \sum_{n=1}^{N'_{I_i}} \frac{1}{K_{I_i}} \sqrt{\frac{K_{I_i} \ln(K_{I_i})}{(e-1)n}} \\ &\approx \int_{n=1}^{N'_{I_i}} \sqrt{\frac{\ln(K_{I_i})}{K_{I_i}(e-1)n}} \\ &\approx 2\sqrt{N'_{I_i}} \sqrt{\frac{\ln(K_{I_i})}{K_{I_i}(e-1)}} \end{aligned} \quad (6.6)$$

with  $N'_{I_i} = \sum_{x'_i} n_{x'_i}$ .

Thus, the visit count corresponding to bidding on set of items  $x_i$  at information set  $I_i$  is reset to:

$$n'_{x_i} \leftarrow \max \left( 0, n_{x_i} - 2\sqrt{N'_{I_i}} \sqrt{\frac{\ln(K_{I_i})}{K_{I_i}(e-1)}} \right) \quad (6.7)$$

Player  $i$  then bids on set of items  $x_i$  with probability  $\frac{n'_{x_i}}{\sum_{x'_i} n'_{x'_i}}$ .

It is important to note that, in  $SMS^\alpha$ , withdrawing visits caused by exploration was unneeded as the final move selection was based on maximising the expected scalarised reward.

## 6.4 Three determinization approaches

As already explained in Section 3.8.5, determinization consists in generating an instance of the equivalent game where all hidden information is assumed known to all players. In this section, we present three different determinization approaches in order to adapt  $SMSS_{EXP3}^\alpha$  to the incomplete information framework of the SAA-inc game. The first approach consists in using the expected value of the opponents' type distribution and apply  $SMSS_{EXP3}^\alpha$  to the resulting determinized SAA-inc game with complete information. The second approach consists in generating many determinized games and apply  $SMSS_{EXP3}^\alpha$  to each different instance. The results obtained for each instance are then combined and a final move is selected. The third approach consists in generating a unique search tree and draw a different determinization at each search iteration. Only the opponents' moves consistent with the drawn determinization are considered, i.e. only the moves within the opponents' drawn budgets are considered.

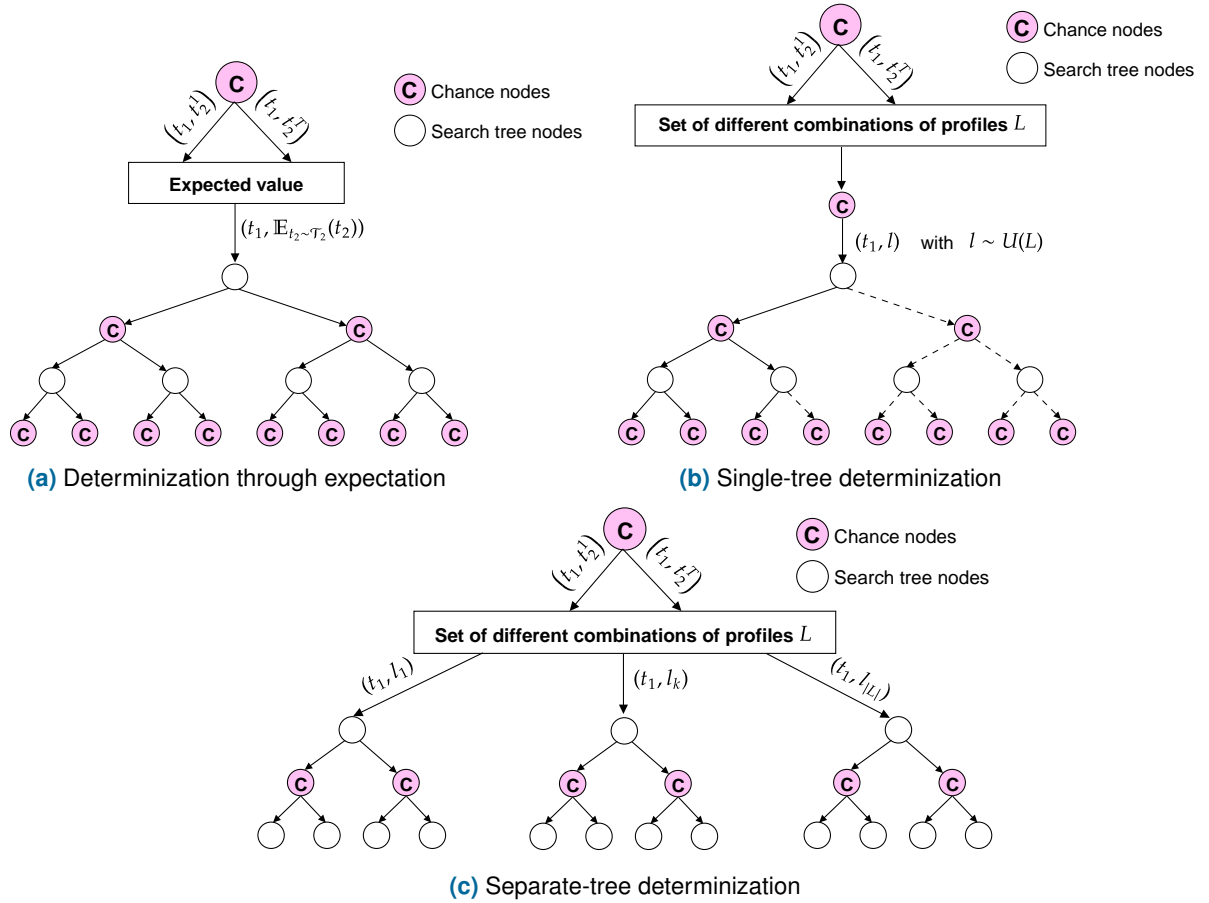
For the two last approaches, it is necessary to pick multiple types for the root player's opponents. Although these could be sampled randomly from the type distribution, given the fact that computing an initial prediction of closing prices for each combination of types is a relatively long process, we only select a few types for each opponent. More precisely, we do not exactly select types but profiles. The selection of these profiles is described in Section 6.5.

**Definition 6.1.** A profile in SAA-inc is a pair  $(v, b)$  where  $v$  corresponds to a value function and  $b$  to a budget. The value function  $v$  only needs to be normalised ( $v(\emptyset) = 0$ ), finite and verify the free disposal condition as described in Section 2.1. The budget  $b$  only needs to be positive and finite.

Thus, a profile  $(v, b)$  does not necessarily belong to the support of a type distribution. However, any type  $(v_i, b_i) \in \text{supp}(\mathcal{T}_i)$  of player  $i$  is a profile.

In Figure 6.2, we represent the three determinization approaches applied to a SAA-inc game with two players from the point of view of player 1. The first chance node in each approach corresponds to the random draw by Nature of the opponent's type, i.e. player 2. The vertices in each search tree correspond to the players' joint actions. For better legibility, we set  $T = |\text{supp}(\mathcal{T}_2)|$ . In Figure 6.2(a), we represent our first simple determinization approach. It computes the expected value of the type distribution of player 2 and then

applies  $SMS_{EXP3}^\alpha$  to the resulting game. In Figure 6.2(b), we represent the single-tree determinization approach. It computes a set of different profiles  $L$  for player 2 and, then, generates a single search tree where, at each search iteration, a profile  $l$  is drawn uniformly from  $L$ . Only moves consistent with profile  $l$  are considered. Vertices with dashed lines represent parts of the search tree which are not consistent with  $l$ . In Figure 6.2(c), we represent the separate-tree determinization approach. It computes the set of different profiles  $L$  for player 2 and then, for each profile, generates a separate search tree on which is applied  $SMS_{EXP3}^\alpha$ .



**Fig. 6.2.:** Representation of the three different determinization approaches applied to a SAA-inc game with two players from the point of view of player 1.

### 6.4.1 Determinization through expectation

A first simple approach to apply  $SMS_{EXP3}^\alpha$  to SAA-inc is to consider the corresponding bidding game with complete information where opponents play according to their expected value of their type distribution. More formally, we suppose that an opponent  $i$  with value

distribution  $\mathcal{F}_i$  and budget distribution  $\mathcal{B}_i$  plays with value function  $\bar{v}_i$  and budget  $\bar{b}_i$  defined hereafter:

$$\begin{cases} \forall X \in \mathcal{P}(\{1, \dots, m\}), \bar{v}_i(X) = \mathbb{E}_{v \sim \mathcal{F}_i}(v(X)) \\ \bar{b}_i = \mathbb{E}_{b \sim \mathcal{B}_i}(b) \end{cases} \quad (6.8)$$

It is important to notice that, as all value functions  $v_i$  sampled from the value distribution  $\mathcal{F}_i$  respect the free disposal condition,  $\bar{v}_i$  respects the free disposal condition and, hence, is a valid value function. However, it is possible that  $(\bar{v}_i, \bar{b}_i) \notin \text{supp}(\mathcal{T}_i)$ . Therefore, the considered bidding game with complete information on which is applied  $SMS_{EXP3}^\alpha$  might not be one of the possible sub-games with complete information of the Bayes-equivalent of the initial SAA-inc game.

*Example.* Let's consider a SAA-inc game with one item and two bidders. The second bidder has one of the four following types with  $b_2^1 \neq b_2^2$  and  $v_2^1 \neq v_2^2$ :  $(v_2^1, b_2^1)$ ,  $(v_2^1, b_2^2)$ ,  $(v_2^2, b_2^1)$  and  $(v_2^2, b_2^2)$ . Thus,  $\bar{v}_2 = \frac{v_2^1 + v_2^2}{2} \notin \text{supp}(\mathcal{F}_i)$  and  $\bar{b}_2 = \frac{b_2^1 + b_2^2}{2} \notin \text{supp}(\mathcal{B}_i)$ .

As we see in the above example, using expected values can lead to a bidding game that cannot appear in the reality given the type distributions. The expected values could even lead to unfeasible combinations of types. Considering this expected game does not pose however an issue in our specific case of auctions. We name this determinization approach  $\overline{SMS_{EXP3}^\alpha}$ . In the case of complete information, playing  $\overline{SMS_{EXP3}^\alpha}$  is equivalent to playing  $SMS_{EXP3}^\alpha$ . The initial prediction of closing prices  $p^*$  is computed the same way as in  $SMS^\alpha$  on the corresponding determinized game.

## 6.4.2 Separate-tree determinization

One of the main issues of computing a bidding strategy by just using the expected value of the opponents' type distributions is that we only use one feature of the distribution. For instance,  $\overline{SMS_{EXP3}^\alpha}$  returns the same mixed policy regardless of the variance of the type distributions. Thus, a simple approach to better exploit the type distribution is to generate separate trees for different determinizations and then combine the results [Cowling, 2012a].

One way of selecting a determinization would be to sample randomly from the opponents' type distribution  $\mathcal{T}_{-i}$ . However, due to computational reasons, we cannot sample many determinizations as computing an initial prediction of closing prices and running  $SMS_{EXP3}^\alpha$  on each separate tree is too long. Thus, as we are forced to select only a few, we decide to



design determinized profiles of opponents which are representative of the type distribution. The selection of these profiles is detailed in Section 6.5.

They are mainly two aspects of the algorithm that we detail here. The first aspect is which actions are expanded at the root node for the concerned player. The second aspect is the final selection move of the overall algorithm given the results obtained in each separate tree.

#### 6.4.2.1 Expansion phase at the root

As expanding all actions at the root of a separate tree prevents in-depth inspection of promising branches, we need to choose a relatively small set of actions to expand. We could expand a different set of actions at the root for the concerned player depending on each determinization (or combination of profiles) and their resulting prediction of closing prices. However, as the aim is to combine the final results obtained in each separate tree to select a final action, it seems fairer to use the same set of actions at the root node for the concerned player (only for the concerned player, not for the opponents) in each separate tree. We select these actions as follows.

Suppose there is a set of different combinations of profiles  $L$ . In each of these combinations, the concerned player has the same profile, i.e. we assume that every player is aware of the concerned player's real value function and budget. For each combination of profiles  $l \in L$ , we compute an initial prediction of closing prices  $p^*(l)$  and generate a separate tree. For the concerned player at the root, we select a maximum of  $N_{act}$  actions. For the same reasons than in  $SMS^\alpha$ , passing its turn is included in the  $N_{act}$  selected actions. The remaining  $N_{act} - 1$  actions correspond to the moves leading to the  $N_{act} - 1$  highest predicted utilities in strategy  $PP$  with initial prediction  $\frac{1}{|L|} \sum_{l \in L} p^*(l)$ . It is important to note that this is not equivalent to selecting the  $N_{act} - 1$  highest average predicted utility according to strategy  $PP$  for each combination of profiles  $l \in L$ . Indeed, if a player  $i$  at information set  $I_i$  is

temporarily winning set of items  $Y_i$  and bids on set of items  $X_i$  with  $P$  the current bid price, we have

$$\begin{aligned}
\frac{1}{|L|} \sum_{l \in L} \sigma_i(Y_i \cup X_i, \rho(p^*(l), P, Y_i)) &= \frac{1}{|L|} \sum_{l \in L} v_i(Y_i \cup X_i) - \sum_{j \in Y_i \cup X_i} \rho_j(p^*(l), P, Y_i) \\
&= v_i(Y_i \cup X_i) - \sum_{j \in Y_i \cup X_i} \sum_{l \in L} \frac{1}{|L|} \rho_j(p^*(l), P, Y_i) \\
&\leq v_i(Y_i \cup X_i) - \sum_{j \in Y_i \cup X_i} \rho_j\left(\sum_{l \in L} \frac{p^*(l)}{|L|}, P, Y_i\right) \\
&= \sigma_i(Y_i \cup X_i, \rho\left(\frac{1}{|L|} \sum_{l \in L} p^*(l), P, Y_i\right))
\end{aligned} \tag{6.9}$$

The term on the left corresponds to the average predicted utility according to strategy  $PP$  over all combinations of profiles while the last term on the right corresponds to the predicted utility in strategy  $PP$  with initial prediction  $\frac{1}{|L|} \sum_{l \in L} p^*(l)$ . The inequality comes from the fact that, for any current bid price  $P$ ,  $p \rightarrow \max(p, P)$  is a convex function. Thus, the perceived price function  $\rho(\cdot, P, Y_i)$  is also convex.

*Remark.* In the specific case where, for each combination of profile  $l \in L$ ,  $p_j^*(l) \geq P_j$  for each item  $j \in Y_i$  and  $p_j^*(l) \geq P_j + \varepsilon$  otherwise, the two approaches are equivalent.

For all other information sets in the separate tree with combination of profiles  $l \in L$ , the expanded set of actions is chosen according to the highest predicted utilities in strategy  $PP$  with initial prediction  $p^*(l)$  as it is the case in  $SMS^\alpha$ .

#### 6.4.2.2 Final move selection

We apply  $SMS_{EXP3}^\alpha$  on each separate tree with the above expansion phase for the root's information set of the concerned player. Thus, for each combination of profiles  $l$ , we obtain a mixed strategy. Our final move selection is based on a variant of majority voting [Nijssen, 2012b]. For each separate tree, an action is drawn from its final mixed strategy. The action which has been drawn the more often (and, thus, has gathered the most votes) across all separate trees is selected as the best move. If two or more moves accumulate the same highest number of votes, then the move with the highest number of visits over all trees amongst the tied moves is selected.

We name the resulting algorithm  $DSMS_{EXP3}^\alpha$  for *Determinized*  $SMS_{EXP3}^\alpha$ . We provide succinct pseudo-code in Algorithm 5. In the case of complete information, the set of different combinations of profiles  $L$  is of size 1. Hence, playing  $DSMS_{EXP3}^\alpha$  is equivalent to playing  $SMS_{EXP3}^\alpha$ .

---

**Algorithm 5**  $DSMS_{EXP3}^\alpha$

---

**Hyperparameters:** The risk-aversion hyperparameter  $\alpha$  and the maximum number of expanded actions  $N_{arms}$

**Inputs computed offline:** A set of combinations of profiles  $L$  and the initial prediction of closing prices  $p^*(l)$  for each combination of profiles  $l \in L$

*// Selecting an expanded action set for the concerned player at the root node*

Selection of the action set  $A_0$  for the concerned player at the root's information set given  $N_{arms}$  and  $\frac{1}{|L|} \sum_{l \in L} p^*(l)$

*// Initialisation*

For each action  $a \in A_0$ , we set the total number of visits  $n_a$  and the total number of votes  $v_a$  to 0

*// Generating and running separate trees*

**for**  $l \in L$  **do**

- Generate a tree with expanded set of actions  $A_0$  at the root's information set for the concerned player
- Run  $SMS_{EXP3}^\alpha$  starting from the initialised tree using determinization  $l$  (the SAA-inc game considered is with complete information)
- When all search iterations have been performed, withdraw the visits due to exploration and update the total number of visits  $n_a$  of each action  $a \in A_0$  with the number of visits  $n'_{a,l}$  obtained in the tree with determinization  $l$ :  $n_a \leftarrow n_a + n'_{a,l}$
- Draw an action  $a \in A_0$  with probability  $\frac{n'_{a,l}}{\sum_{a' \in A_0} n'_{a',l}}$  and update its total number of votes:  $v_a \leftarrow v_a + 1$

**end for**

*// Final selection move*

**return** Action  $a \in A_0$  with the maximal number of votes  $v_a$  or, in case of ties, with the maximal number of visits  $n_a$  amongst the tied moves

---

### 6.4.3 Single-tree determinization

In  $DSMS_{EXP3}^\alpha$ , we generate  $|L|$  separate trees depending on the size of the set of combinations of profiles  $L$ . Another approach is to consider only one tree and to share the

information between different combinations of profiles. More precisely, at each search iteration, a combination of profiles  $l$  is drawn and only the opponents' moves consistent with  $l$  are considered. Single-tree determinization [Pepels, 2016] is also named Single-Observer Information Set MCTS (SO-ISMCTS) [Cowling, 2012a] as nodes in the tree correspond to information sets from the root player's point of view. Indeed, we assume that all players know the real value function and budget of the concerned player. To combine different determinizations in a same tree, we need to modify the selection phase, expansion phase, rollout phase and backpropagation phase of  $SMS_{EXP3}^\alpha$  accordingly.

### 6.4.3.1 Selection

Depending on the combination of profiles  $l \in L$ , the set of actions available at a particular node might vary. Indeed, only opponents' moves consistent with  $l$  are considered, i.e. only moves withing their determinized budget. This issue can be seen as a multi-armed bandit problem where only a subset of the arms is available at each trial. This problem is referred to as *subset-armed bandit* [Cowling, 2012a]. This problem is easily solved for greedy selection methods such as UCT where the selected moves only depends on the current statistics. Indeed, each UCT value can be computed independently for each arm. It only depends on the number of trials where an arm could have been pulled, the number of times an arm was pulled and its sum of rewards. However, the EXP3 algorithm returns a mixed policy by mixing two distributions that both depend on the statistics of all arms considered. Thus, extending EXP3 to the subset-armed bandit problem seems less straightforward. We propose the following modifications.

During the selection phase, each opponent  $i$  with profile  $l_i$  chooses to bid on the set of items  $x_i$  with probability  $\mathbf{P}_i(x_i)$  at information state  $I_i$ :

$$\mathbf{P}_i(x_i) = \frac{\gamma_{I_i}^{l_i}}{K_{I_i}^{l_i}} + \frac{(1 - \gamma_{I_i}^{l_i})}{\sum_{x'_i} e^{\eta_{I_i}^{l_i} (\frac{n_{I_i}^{l_i}}{N_{x'_i}} s_{x'_i}^{\alpha} - \frac{n_{I_i}^{l_i}}{N_{x_i}} s_{x_i}^{\alpha})}} \quad (6.10)$$

where  $K_{I_i}^{l_i}$  corresponds to the number of legal moves which can be played by player  $i$  at  $I_i$  with profile  $l_i$ ,  $n_{I_i}$  the total number of times  $I_i$  has been visited and  $N_{x_i}$  the total number of

times a player  $i$  could have bidden on  $x_i$  at  $I_i$  (number of times this move was considered legal). Both hyperparameters are modified accordingly:

$$\begin{cases} \gamma_{I_i}^{l_i} = \min(1, \sqrt{\frac{K_{I_i}^{l_i} \ln(K_{I_i}^{l_i})}{(e-1)n_{I_i}}} \\ \eta_{I_i}^{l_i} = \frac{\gamma_{I_i}^{l_i}}{K_{I_i}^{l_i}} \end{cases} \quad (6.11)$$

As  $s_{x_i}^\alpha$  acts as an estimate of the sum of scalarised rewards after bidding on  $x_i$  at  $I_i$  over all search iterations where  $I_i$  is encountered in the selection phase and bidding on  $x_i$  is legal, then  $\frac{n_{I_i}}{N_{x_i}} s_{x_i}^\alpha$  acts as an estimate over all search iterations where  $I_i$  is encountered. This enables fair comparison between the different moves in the Gibbs distribution. Without this rescale, moves which can not be played with low budget profiles would be disfavoured when high budget profiles are drawn. The following example highlights the importance of this rescale.

*Example.* Suppose that a player has the choice between 2 moves:  $a_1$  and  $a_2$ . The player can either have profile  $l_1$  or  $l_2$ . Profiles are drawn randomly. In profile  $l_1$ , only action  $a_1$  is legal whereas, in profile  $l_2$ , both moves are legal. Both moves return the same deterministic reward  $r = 1$ . For each trial, we first draw a profile  $l_i$ , then an action is drawn from the conditional probability  $\mathbf{P}(\cdot|l_i)$  and we receive a reward  $r$ . Given the fact that both moves lead to the same reward  $r$ , they should have the same expected probability of getting selected when  $l_2$  is drawn. However, the estimate  $s_{a_1}^{3N}$  of the sum of rewards of bidding on  $a_1$  after  $3N$  trials theoretically gives in expectation

$$\begin{aligned} \mathbb{E}(s_{a_1}^{3N}) &= 3N(\mathbf{P}(l_1)\mathbf{P}(a_1|l_1)\frac{r}{\mathbf{P}(a_1|l_1)} + \mathbf{P}(l_2)\mathbf{P}(a_1|l_2)\frac{r}{\mathbf{P}(a_1|l_2)}) \\ &= 3N(0.5 \times 1 \times \frac{1}{1} + 0.5 \times 0.5 \times \frac{1}{0.5}) \\ &= 3N \end{aligned} \quad (6.12)$$

while the estimate  $s_{a_2}^{3N}$  of the sum of rewards of bidding on  $a_2$  after  $3N$  trials gives

$$\begin{aligned} \mathbb{E}(s_{a_2}^{3N}) &= 3N(\mathbf{P}(l_2)\mathbf{P}(a_2|l_2)\frac{r}{\mathbf{P}(a_2|l_2)}) \\ &= 3N(0.5 \times 0.5 \times \frac{1}{0.5}) \\ &= \frac{3}{2}N \end{aligned} \quad (6.13)$$

Thus, if no rescale is done, when  $l_2$  is drawn,  $a_1$  has far more chances of getting selected by the EXP3 algorithm than  $a_2$ . Moreover, the probability of selecting  $a_2$  when  $l_2$  is drawn converges to 0. However, by using our rescale, we have that  $\frac{1}{\mathbb{P}(a_1 \text{ is legal})} \mathbb{E}(s_{a_1}^{3N}) = \mathbb{E}(s_{a_1}^{3N}) = 2 \times \mathbb{E}(s_{a_2}^{3N}) = \frac{1}{\mathbb{P}(a_2 \text{ is legal})} \mathbb{E}(s_{a_2}^{3N})$  and, thus, both moves have the same expected probability of getting selected by our modified EXP3 algorithm when  $l_2$  is drawn.

At the end of each selection step, for all sets of items  $x_i$  of  $I_i$  within the budget of player  $i$  according to  $l_i$ , we set  $N_{x_i} \leftarrow N_{x_i} + 1$ .

One slight issue with this approach is that the uniform distribution decreases in  $\frac{1}{\sqrt{n_{I_i}}}$  regardless of profile  $l_i$ . Thus, moves which can not be played with low budget profiles have less chance of being explored than the others. Nevertheless, given the fact that the number of different profiles  $l_i$  for an opponent  $i$  is low, this should not be too problematic.

During the selection phase, we apply the above modified EXP3 algorithm to the information sets of the concerned player's opponents. For the concerned player, the usual EXP3 algorithm is used as selection strategy as its profile never changes during the search iterations.

#### 6.4.3.2 Expansion phase

As it is the case for  $SM S_{EXP3}^\alpha$ , only a maximum number of actions  $N_{arm}$  can be expanded per information set. Passing its turn is always included in the  $N_{arm}$  actions. At each search iteration, a combination of profiles  $l$  is drawn. For each player  $i$  at information state  $I_i$  temporarily winning set of items  $Y_i$  with eligibility  $e_i$ , the action of bidding on set of items  $x_i$  is expanded if  $\sigma_i^\alpha(Y_i \cup x_i, \rho(p^*(l), P, Y_i))$  has the highest value between the remaining legal unexpanded actions. Only sets of items  $x_i$  which respect the eligibility and budget constraints given by  $e_i$  and profile  $l_i$  are considered. Statistics for each action are initialised as follows:

- \*  $s_{x_i}^\alpha \leftarrow 0$

- \*  $n_{x_i} \leftarrow 0$

- \*  $N_{x_i} \leftarrow 1$

Unlike  $SM S_{EXP3}^\alpha$  where there is a clear division between the selection phase and the expansion phase, this is not the case here. Both are intertwined. This is due to the fact that actions are added one by one in each information set and depend on the combination of profiles drawn. We can divide these in two separate phases:

- **Selection of expanded nodes:** If the node is already expanded, i.e. a same configuration of temporary winners, bid prices and eligibilities has already been added to the search tree, then we check for each information state  $I_i$  the two following cases:
  - If the maximum number of actions  $N_{arm}$  has already been reached or all legal actions have already been added to the search tree, then we apply the selection strategy described in the last subsection and draw an action from the corresponding mixed policy.
  - Otherwise, a new action is added to the information set as described above.

Each player then plays its corresponding action. This process continues until a non-expanded node is reached.

- **Adding a non-expanded node:** If a non-expanded node is reached, then, for each information set  $I_i$ , an action is expanded as detailed above and  $N_{I_i} \leftarrow 0$ .

#### 6.4.3.3 Rollout

Similarly to  $SM S_{EXP3}^\alpha$ , at the beginning of each rollout phase, we set  $p_i^* = p^*(l) + \eta_i$  with  $\eta_i \sim U([- \varepsilon, \varepsilon]^m)$  and  $l$  the combination of profiles drawn at this search iteration. Each player  $i$  then plays  $PP$  with initial prediction of closing prices  $p_i^*$  during the entire rollout.

#### 6.4.3.4 Backpropagation

Let  $V_i^\alpha$  be the scalarised utility obtained by player  $i$  at the end of the rollout. Let  $x_i$  be the set of items on which player  $i$  bid at information state  $I_i$  for one of the selected nodes. The statistics stored for  $I_i$  are updated as follows:

$$* s_{x_i}^\alpha \leftarrow s_{x_i}^\alpha + \frac{V_i^\alpha}{\mathbf{P}_i(x_i)}$$

$$* n_{x_i} \leftarrow n_{x_i} + 1$$

\*  $N_{I_i} \leftarrow N_{I_i} + 1$

In the specific case where it is the first time that  $x_i$  is selected ( $n_{x_i} = 0$  before the backpropagation phase), we set  $s_{x_i}^\alpha \leftarrow V_i^\alpha$ .

We name the resulting algorithm  $SDSMS_{EXP3}^\alpha$  for *Single-tree Determinization*  $SMS_{EXP3}^\alpha$ . We provide succinct pseudo-code in Algorithm 6. It uses the same transposition table as  $SMS^\alpha$  described in Section 5.4.7. It is important to note that, in case of complete information, playing  $SDSMS_{EXP3}^\alpha$  is equivalent to playing  $SMS_{EXP3}^\alpha$ .

---

**Algorithm 6**  $SDSMS_{EXP3}^\alpha$

---

**Hyperparameters:** The risk-aversion hyperparameter  $\alpha$  and the maximum number of expanded actions  $N_{arms}$

**Inputs computed offline:** A set of combinations of profiles  $L$  and the initial prediction of closing prices  $p^*(l)$  for each combination of profiles  $l \in L$

**for**  $t$  search iterations **do**

    // *Select combination of profiles*

$l \sim U(L)$  where  $U$  is the uniform distribution

    // *Selection of expanded nodes*

    Select a path consistent with  $l$  until a non-expanded node is reached using the *Selection of expanded nodes* process with  $N_{arms}$  and  $p^*(l)$  described in Section 6.4.3.2

    // *Adding a non-expanded node*

    Add the non-expanded node using  $l$  and  $p^*(l)$

    // *Rollout*

    Simulate a SAA-inc game where each bidder  $i$  plays a PP strategy with initial prediction of closing prices  $p_i^* = p^*(l) + \eta_i$  ( $\eta_i \sim U([- \varepsilon, \varepsilon]^m)$ )

    // *Backpropagation*

    Update statistics of the selected actions using the results of the rollout phase

**end for**

    // *Final selection move*

    Withdraw the visits due to exploration and then compute the mixed policy for the concerned player  $i$  at the root's information state  $I_i$  by setting the probability  $\mathbb{P}_i(x_i)$  of bidding on set of items  $x_i$  to  $\frac{n_{x_i}}{\sum_{x'_i} n_{x'_i}}$ .

**return** An action drawn from the mixed policy

---



## 6.5 Generating combinations of profiles

Due to time constraints, we can only sample a few determinizations as computing an initial prediction of closing prices or running a separate tree for each determinization is relatively long (a few minutes). One approach would be to sample them randomly from the type distribution. However, given the restrictive number of determinizations we allow ourselves, there is fair chance that the sampling is not representative of the opponents' type distributions. As the performance of  $SDSMS_{EXP3}^\alpha$  and  $DSMS_{EXP3}^\alpha$  highly rely on the quality of the sampling, this could lead to poor results. Thus, we decide to design profiles of opponents which are representative of the type distribution.

Each profile is parametrised by a unique parameter  $\delta$ . An opponent  $i$  with profile  $\delta$  plays with value function  $v_i^\delta$  and budget  $b_i^\delta$  defined as:

$$\begin{cases} \forall X \in \mathcal{P}(\{1, \dots, m\}), v_i^\delta(X) = \mathbb{E}_{v \sim \mathcal{F}_i}(v(X)) + \delta \sqrt{\mathbb{V}_{v \sim \mathcal{F}_i}(v(X))} \\ b_i^\delta = \mathbb{E}_{b \sim \mathcal{B}_i}(b) + \delta \sqrt{\mathbb{V}_{b \sim \mathcal{B}_i}(b)} \end{cases} \quad (6.14)$$

It is plausible that a profile  $(v_i^\delta, b_i^\delta) \notin \text{supp}(\mathcal{T}_i)$ . Nevertheless, we believe that by selecting different symmetric profiles, i.e. if profile  $\delta$  is selected then profile  $-\delta$  is selected, our sampling method is far more representative of the type distributions than just sampling randomly a few determinizations from the type distribution.

*Example.* Consider an SAA-inc game between 4 players. A possible combination of profiles is to assign to each opponent a profile  $(v_i^\delta, b_i^\delta)$  with  $\delta \in \{-1, 0, 1\}$ . Thus, each opponent can have one of the three profiles. Therefore, the set of combinations of profiles  $L$  has 27 elements.

The higher  $\delta$  is, the higher are the value function and budget of an opponent playing with profile  $\delta$  and, thus, the more aggressive this opponent will be in the determinized game. If only strong profiles (high value functions and high budgets) are sampled, then both algorithms  $SDSMS_{EXP3}^\alpha$  and  $DSMS_{EXP3}^\alpha$  are more risk-averse as they try to reduce the risk of exposure and the own price effect. Symmetrically, the lower  $\delta$  is, the more risk-averse the opponent will be in the determinized game. If only weak profiles (low value functions and low budgets) are sampled, then both algorithms  $SDSMS_{EXP3}^\alpha$  and  $DSMS_{EXP3}^\alpha$  bid more aggressively as they take advantage of the weak position of their opponents. Thus, to represent appropriately one's type distribution, we select symmetric values  $\delta$  and  $-\delta$  in order to obtain as many strong profiles as weak profiles for each opponent.

## 6.6 Tracking bid exposure

Until now, we have never used the history of past moves to enhance our algorithms. We have only always focused on the current information set. Past bids submitted by a bidder can be seen as an important source of information as one might give away information about its preferences or budget. This could be used to narrow down an opponent's possible types. Ideally, one would like to end up in a game with complete information with a correct estimation of all its opponents' types. However, such inference is complicated in practice as a bidder could eventually use bluff in order to fool others if it knew its moves are being tracked. Moreover, even though we believe that inferring a bidder's value function from its past moves given a certain prior is possible, this remains a very difficult task and is not considered in this thesis.

We propose an easy inference method to update one's belief about its opponents' budget distribution in a SAA-inc game. It is based on tracking one's bid exposure [Bulow, 2009], i.e. the sum of all bids placed by a bidder during a given round, including its temporary winning bids from the prior round. As bidders are not authorised to bid over their budget in this thesis, the bid exposure is a perfect estimate of the lower bound of one's budget. Therefore, by tracking an opponent's bid exposure, one can update its belief of its opponent's budget distribution and, thus, narrow down its types. This is described formally in Definition 6.1.

**Definition 6.1. Budget inference through tracking bid exposure:** Let  $\mathcal{B}_i$  be the initial budget distribution of player  $i$  and  $\hat{\mathcal{B}}_i$  be the inferred budget distribution of player  $i$  using its bid exposure  $\hat{b}_i$ . Let  $\mathbf{P}_{\mathcal{B}_i}$  be the probability associated to  $\mathcal{B}_i$  and  $\mathbf{P}_{\hat{\mathcal{B}}_i}$  be the probability associated to  $\hat{\mathcal{B}}_i$ . Thus, for all  $b \in \text{supp}(\mathcal{B}_i)$ ,  $\hat{\mathcal{B}}_i$  is defined as follows:

$$\left\{ \begin{array}{l} \text{If } b < \hat{b}_i, \mathbf{P}_{\hat{\mathcal{B}}_i}(b) = 0 \\ \text{Otherwise, } \mathbf{P}_{\hat{\mathcal{B}}_i}(b) = \frac{\mathbf{P}_{\mathcal{B}_i}(b)}{\sum_{\substack{b' \in \text{supp}(\mathcal{B}_i) \\ b' \geq \hat{b}_i}} \mathbf{P}_{\mathcal{B}_i}(b')} \end{array} \right. \quad (6.15)$$

Thus, we replace  $\mathcal{B}_i$  by  $\hat{\mathcal{B}}_i$  in  $\overline{SMS}_{EXP3}^\alpha$ ,  $DSMS_{EXP3}^\alpha$  and  $SDSMS_{EXP3}^\alpha$ . This enables us to have more precise estimate of our opponents' budget distribution. All profiles in  $DSMS_{EXP3}^\alpha$  and  $SDSMS_{EXP3}^\alpha$  are also generated by using  $\hat{\mathcal{B}}_i$  instead of  $\mathcal{B}_i$ . Thus, some profiles might change throughout the auction. Ideally, one should recompute the initial prediction of closing prices  $p^*(l)$  each time a combination of profiles  $l$  changes. However,

as the predictions are computed offline as it is time-consuming, we conserve the same initial prediction of closing price for a combination of profiles even if the determinized budget of an opponent is slightly modified.

## 6.7 Type generation

In this section, we describe a general framework to generate type distributions for an SAA-inc bidding game with a predetermined level of uncertainty. As we assume throughout this thesis that one's budget is chosen independently from its value function, budget distributions and value distributions can be generated separately. Thus, we first present a general framework for generating value functions and, then, we present a general framework for generating budgets.

For simplicity, we present hereafter each framework with continuous distributions. In practice, once the value distributions or budget distributions obtained, we discretize them with a small step (a step of  $10^{-2}$  is chosen in our numerical experiments).

### 6.7.1 Value generation

Most works in the literature consider that the value distribution  $\mathcal{F}_i$  of player  $i$  is Uniform or Log-Normal [Nedelec, 2022; Yuan, 2014]. For instance, Goeree et al. [Goeree, 2003] use Uniform distributions to model the values and costs of each bidder in single-object first-price, second-price or English auctions. However, very few works have considered value distributions for multi-item auctions, especially in SAA. One exception is Reeves et al [Reeves, 2005] which uses an SAA for a scheduling problem. They initialise each slot with a value drawn uniformly from  $[0, 50]$ . Then, they apply pruning on the slots that violate monotonicity in order to obtain super-additive value functions. This setting is also used by Wellman et al [Wellman, 2008]. In this subsection, we propose a new framework to generate general value distributions which extends Settings 1 and 2 to incomplete information.

### 6.7.1.1 Complementarity distributions

In order to build the value distribution  $\mathcal{F}_i$  of player  $i$ , we introduce a new distribution  $\mathcal{G}_i^{\eta_v}$  that we name complementarity distribution. It is based on a parameter  $\eta_v \in [0, 1]$  which determines the level of certainty that bidders have on the value function of their opponents. In order to match with Settings 1 and 2, we consider two disjoint cases when building  $\mathcal{G}_i^{\eta_v}$  depending on whether  $X$  is a singleton or not. The formal construction of  $\mathcal{G}_i^{\eta_v}$  is described below:

**Definition 6.1.** *Let  $\Gamma$  be an instance of an SAA-inc game with  $n$  bidders,  $m$  items, a bid increment  $\varepsilon$  and with a level of certainty on the values of  $\eta_v \in [0, 1]$ . Let  $V$  be the maximum surplus of complementarity gained by obtaining an extra item (as defined in Settings 1 and 2). For each player  $i$ , the complementarity distribution  $\mathcal{G}_i^{\eta_v}$  is built as follows:*

- *If the set of items  $X$  is a singleton, then we draw  $c_X \sim U([0, \eta_v V])$  and define  $\mathcal{G}_{i,X}^{\eta_v} = U([c_X, c_X + (1 - \eta_v)V])$ . Hence, the size of the support of  $\mathcal{G}_{i,X}^{\eta_v}$  is  $(1 - \eta_v)V$ .*
- *If the set of items  $X$  is not a singleton, then we draw  $c_X \sim U([0, 2\eta_v V])$  and define  $\mathcal{G}_{i,X}^{\eta_v} = U([c_X, c_X + 2(1 - \eta_v)V])$ . Hence, the size of the support of  $\mathcal{G}_{i,X}^{\eta_v}$  is  $2(1 - \eta_v)V$ .*

where  $U$  is the uniform distribution.

We sample  $c_X$  uniformly from  $[0, \eta_v V]$  or  $[0, 2\eta_v V]$ , depending on whether or not the set of items  $X$  is a singleton, so that our framework extends Settings 1 and 2 to incomplete information. If  $\eta_v = 1$ , then the complementarity distributions are dirac delta distributions and the value functions are common knowledge. If  $\eta_v = 0$ , then for every bidder  $i$ ,  $\mathcal{G}_{i,X}^{\eta_v} = U([0, V])$  if  $X$  is a singleton and  $\mathcal{G}_{i,X}^{\eta_v} = U([0, 2V])$  otherwise.

### 6.7.1.2 Value distributions

From the complementarity distributions, we are able to draw the private value functions through the following process.

**Setting 3.** Let  $\Gamma$  be an instance of an SAA-inc game with  $n$  bidders,  $m$  items, a bid increment  $\varepsilon$  and with a level of certainty on the values of  $\eta_v \in [0, 1]$ . Let  $\mathcal{G}_i^{\eta_v}$  be the complementarity distribution of player  $i$ .  $\mathcal{G}_i^{\eta_v}$  is common knowledge. From this known distribution, we draw a unique vector of size  $2^m$ ,  $C_i \sim \mathcal{G}_i^{\eta_v}$ , which is then used to build the private valuation  $v_i$  as follows:

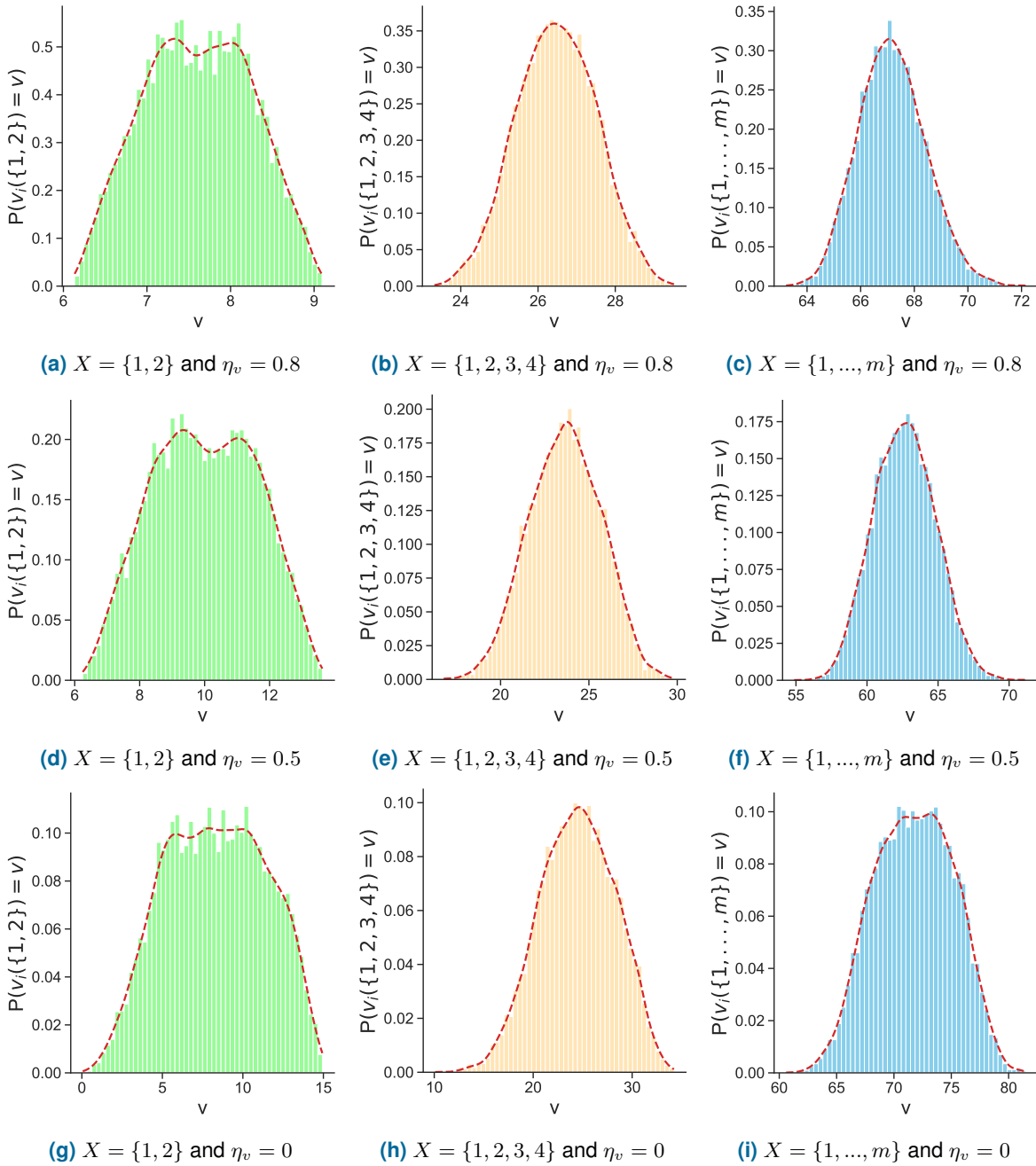
$$v_i(X) = \max_{j \in X} v_i(X \setminus \{j\}) + C_{i,X} \quad (6.16)$$

Similarly to the private value functions, the value distribution  $\mathcal{F}_i$  is obtained empirically through the same process for every player  $i$ . Equation 6.16 ensures that all value distributions  $v_i \sim \mathcal{F}_i$  respect the free disposal condition. Moreover,  $C_{i,X} \sim \mathcal{G}_{i,X}^{\eta_v}$  can be interpreted as the complementarity surplus obtained when purchasing a set of items  $X$  instead of purchasing the set  $X \setminus \{j\}$  (with  $j \in X$ ) which maximises  $v_i$ . In Settings 1 and 2, this complementarity surplus was in  $[0, V]$  if  $X$  was a singleton and in  $[0, 2V]$  otherwise. This explains our disjunction of cases when generating the complementarity distribution  $\mathcal{G}_i^{\eta_v}$ .

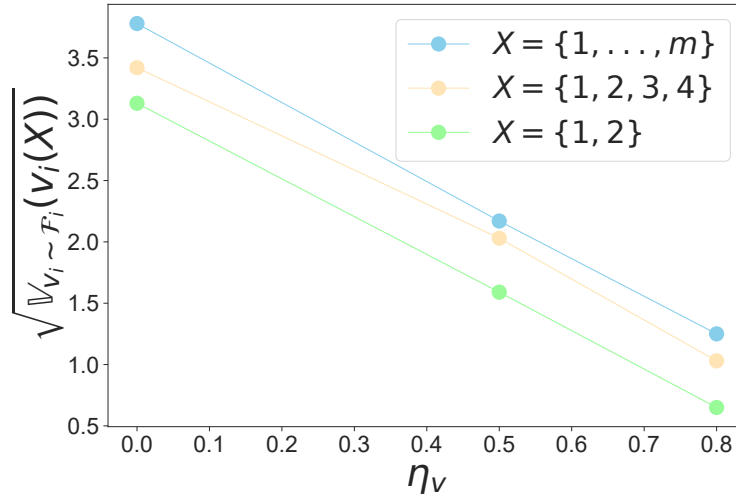
### 6.7.1.3 Numerical examples

Although it is relatively straightforward how the level of certainty on the values  $\eta_v$  affects the complementarity distribution  $\mathcal{G}_i^{\eta_v}$  of a player  $i$ , it is less obvious how it impacts its value distribution  $\mathcal{F}_i$ . Thus, we represent in Figure 6.3, the empirical value distribution of obtaining different sets of items  $X$  for different levels of certainty by drawing thousands of  $v_i \sim \mathcal{F}_i$ . Whereas the complementarity distributions are based on uniform distributions, we notice that the empirical value distribution of obtaining a set of items  $X$  seems to approach a Gaussian distribution through our process when  $|X|$  increases. Moreover, as expected, the support of the empirical distributions widens when  $\eta_v$  decreases. For instance, for  $X = \{1, \dots, m\}$ , the size of the support of the empirical distribution is approximately 8 when  $\eta_v = 0.8$  and 20 when  $\eta_v = 0$ .

We represent in Figure 6.4 the evolution of the standard deviation for each empirical value distribution for obtaining a set of items  $X$  versus  $\eta_v$ . We notice that the standard deviation of the empirical distribution increases linearly when  $\eta_v$  decreases.



**Fig. 6.3.:** Empirical value distributions for obtaining sets  $\{1, 2\}$ ,  $\{1, 2, 3, 4\}$  and  $\{1, \dots, m\}$  for level of certainty  $\eta_v \in \{0, 0.5, 0.8\}$  by sampling  $10^4$  value functions  $v_i$  from the value distribution  $\mathcal{F}_i$  of player  $i$ . The complementarity distribution  $\mathcal{G}_i^{\eta_v}$  and value distribution  $\mathcal{F}_i$  are built with  $V = 5$  and  $m = 9$ .



**Fig. 6.4.:** Evolution of the standard deviation of the empirical value distributions for obtaining sets  $\{1, 2\}$ ,  $\{1, 2, 3, 4\}$  and  $\{1, \dots, m\}$  depending on the level of certainty  $\eta_v$ . We sample  $10^4$  value functions  $v_i$  from the value distribution  $\mathcal{F}_i$  of player  $i$ . The complementarity distribution  $\mathcal{G}_i^{\eta_v}$  and value distribution  $\mathcal{F}_i$  are built with  $V = 5$  and  $m = 9$ .

## 6.7.2 Budget generation

For the generation of budget distributions, we use a similar process than for the generation of value distributions. We extend the generation of budgets in Setting 2 to incomplete information.

**Setting 4.** Let  $\Gamma$  be an instance of an SAA-inc game with  $n$  bidders,  $m$  items, a bid increment  $\varepsilon$  and with a level of certainty on the budgets of  $\eta_b \in [0, 1]$ . For each player  $i$ , we define the size of the support of their budget distribution as  $d_{\eta_b} = (1 - \eta_b)(b_{max} - b_{min})$  where  $b_{min}$  and  $b_{max}$  are respectively the minimal and maximal budget that a player can have. The budget distribution  $\mathcal{B}_i$  is built as follows:

- We draw  $B_i \sim U([b_{min}, b_{max} - d_{\eta_b}])$
- We define  $\mathcal{B}_i = U([B_i, B_i + d_{\eta_b}])$

where  $U$  is the uniform distribution. The private budget  $b_i$  of bidder  $i$  is then drawn from  $\mathcal{B}_i$ .

If  $\eta_b = 1$ , then the budgets are common knowledge. Moreover if  $\eta_v = 1$ , then the induced SAA-inc bidding game is with complete information. If  $\eta_b = 0$ , then a bidder only knows that the budget of its opponents is between  $b_{min}$  and  $b_{max}$ .

## 6.8 Extensive experiments

### 6.8.1 Benchmark

In this section, we measure the performance of our three determinization approaches for different levels of certainty by notably comparing them to the bidding strategies developed in Section 2.4. More precisely, we consider the following 8 bidding algorithms:

- $\overline{SMS}_{EXP3}^\alpha$  described in Section 6.4.1 which uses the expected value of the opponents' type distributions and applies  $SMS_{EXP3}^\alpha$  to the resulting game with complete information.
- $DSMS_{EXP3}^\alpha$  described in Section 6.4.2 which generates a separate determinized tree for each combination of profiles on which  $SMS_{EXP3}^\alpha$  is applied. The final move is selected through a variant of majority voting.
- $SDSMS_{EXP3}^\alpha$  described in Section 6.4.3 which is a single-tree determinization approach which uses a variant of  $SMS_{EXP3}^\alpha$  to share the information between different combinations of profiles. At each search iteration, a combination of profiles is drawn and only the opponents' moves consistent with this combination of profiles are considered.
- $CSMS_{EXP3}^\alpha$  (or *Cheating*  $SMS_{EXP3}^\alpha$ ) is a direct application of  $SMS_{EXP3}^\alpha$  on the SAA-inc game with complete information and is mainly used as benchmark to highlight the impact of uncertainty on our three determinization approaches. It "cheats" by observing the real state of the game before submitting a bid. Hence, an  $CSMS_{EXP3}^\alpha$  bidder has no uncertainty regarding the types of its opponents.
- Strategy *SB* [Milgrom, 2000] described in Section 2.4.1.
- The *EPE* algorithm [Wellman, 2008] described in Section 2.4.2.2.
- The *EDPE* algorithm [Wellman, 2008] described in Section 2.4.2.3.
- The *SCPD* algorithm [Wellman, 2008] described in Section 2.4.3.1.



As it was done in the last chapter, we extend the four strategies EDPE, EPE, SCPD and SB (equivalent to  $PP$  with  $P^{init} = 0$ ) to budget and eligibility constrained environments in the same way as it is done in Definition 5.1. We denote by  $S_{SMS_{EXP_3}^\alpha} = \{SDSMS_{EXP_3}^\alpha, DSMS_{EXP_3}^\alpha, \overline{SMS_{EXP_3}^\alpha}, CSMS_{EXP_3}^\alpha\}$  the set of approaches based on  $SMS_{EXP_3}^\alpha$  and by  $S_{PPB} = \{SB, EPE, EDPE, SCPD\}$  the set of perceived-price bidding strategies (that we name  $PPB$  from now on). In all experiments, none of the bidders are aware of their opponents' strategy.

We first describe our experimental setting by presenting the instances on which the above strategies are confronted and the specificities of each algorithm such as the generation of combinations of profiles or hyperparameters for our determinization approaches. We then compare the different approaches.

As done in the last chapter, our analysis is divided into three parts: expected utility, own price effect and exposure. Each part is divided into three portions. The first portion consists in showing that our determinization approach significantly outperforms any strategy  $B \in S_{PPB}$ . The second portion consists in comparing our three determinization approaches. The last portion consists in analysing the impact of uncertainty on the considered performance indicators for our three determinization approaches. For each portion, we will generally consider two types of scenarios: either bidders play different strategies or they all play the same strategy.

As done in the last chapter, measures on performance indicators such as the average price paid per item won are obtained by confronting a strategy  $A$  to a strategy  $B$  and averaging the results over the two possible strategy profiles:  $(A, B, B)$  and  $(A, A, B)$ . For instance, if  $A = SDSMS_{EXP_3}^\alpha$  and  $B = SB$ , the average price paid per item won by  $SDSMS_{EXP_3}^\alpha$  in these two strategy profiles is respectively: 5.88 and 4.94. Hence, the average price payer per item won by  $SDSMS_{EXP_3}^\alpha$  against  $SB$  is 5.41.

## 6.8.2 Experimental setting

We study SAA-inc instances with  $n = 3$ ,  $m = 9$  and  $\varepsilon = 1$ . Each experimental result has been run on 1000 different SAA-inc instances. Value distributions and budget distributions are generated separately through setting 3 and setting 4 with  $V = 5$ ,  $b_{min} = 10$  and  $b_{max} = 40$ . We consider three levels of certainty with  $\eta_v \in \{0, 0.5, 0.8\}$  and  $\eta_b = \eta_v$ . During our analysis, we will mainly focus on  $\eta_v = 0.5$ . Moreover, we never merge results between different levels of certainty. A recap of the features used for the different instances in our

extensive experiments is given in Table 6.1. Considering vastly different levels of certainty (from no information to a relatively precise estimation of opponents' value functions and budgets) enables us to highlight the impact of uncertainty on our three determinization approaches. All experiments are run on a server consisting of Intel®Xeon®E5-2699 v4 2.2GHz processors.

SAA-inc instance features	Notation	Value
Number of players	$n$	3
Number of objects	$m$	9
Bid increment	$\varepsilon$	1
Maximum surplus of complementarity	$V$	5
Level of certainty on the values	$\eta_v$	$\eta_v \in \{0, 0.5, 0.8\}$
Minimal budget per bidder	$b_{min}$	10
Maximal budget per bidder	$b_{max}$	40
Level of certainty on the budgets	$\eta_b$	$\eta_b = \eta_v$

**Tab. 6.1.:** Summary of the features of the SAA-inc instances employed for our extensive experiments

In the two following subsections, we specify how the combinations of profiles are generated for  $DSMS_{EXP3}^\alpha$  and  $SDSMS_{EXP3}^\alpha$  as well as our choice of hyperparameters. It is important to note that we have chosen different risk-aversion hyperparameters  $\alpha$  based on the level of certainty.

### 6.8.2.1 Generating combinations of profiles

$DSMS_{EXP3}^\alpha$  and  $SDSMS_{EXP3}^\alpha$  both rely on a set of combinations of profiles  $L$  in order to choose on which set of items to bid. Even though it is done offline, computing an initial prediction of final prices for each combination of profiles is time-consuming. Thus, we decide to assign to each opponent  $i$  a profile  $(v_i^\delta, b_i^\delta)$  with  $\delta \in \{-1, 0, 1\}$ . Hence, as there are only two opponents, the number of different combinations of profiles in  $L$  is 9.

### 6.8.2.2 Hyperparameters and thinking time

For each different level of certainty, we select a different hyperparameter  $\alpha$  for each approach based on  $SMS_{EXP3}^\alpha$  that we display in Table 6.2. These are selected through grid-search (more details are given in the Remark of Section 6.8.3.1).

Moreover, the maximum number of expanded actions per information set  $N_{act}$  of all MCTS approach except  $DSMS_{EXP3}^\alpha$  is set to 20. For  $DSMS_{EXP3}^\alpha$ ,  $N_{act}$  is set to 10. This is

$\eta_v$	0	0.5	0.8
$CSMS_{EXP3}^\alpha$	0.8	0.8	0.8
$SMS_{EXP3}^\alpha$	0.2	0.8	0.8
$DSMS_{EXP3}^\alpha$	0.1	0.2	0.2
$SDSMS_{EXP3}^\alpha$	0.7	0.5	0

**Tab. 6.2.:** Hyperparameter  $\alpha$  used for each approach based on  $SMS_{EXP3}^\alpha$  and each level of certainty  $\eta_v = \eta_b$  in a SAA-inc game ( $n = 3, m = 9, \varepsilon = 1$ )

due to the fact that  $DSMS_{EXP3}^\alpha$  generates 9 separate trees instead of one for the other approaches and, thus, in order to obtain satisfactory results given thinking time constraints, it must create smaller search trees.

Each algorithm is given respectively 100 seconds of thinking time except  $DSMS_{EXP3}^\alpha$  where we give 50 seconds per tree. Initial prediction of closing prices for each algorithm are done offline.

### 6.8.3 Expected utility

To analyse a strategy's expected utility, we apply the same pairwise comparison as in Section 5.5.3.1 by considering the normal form game in expected utility where each player has the choice between playing either a strategy  $A$  or a strategy  $B$ . We map each strategy profile to the estimated expected utility obtained by each player in 1000 SAA-inc instances.

#### 6.8.3.1 Comparing approaches based on $SMS_{EXP3}^\alpha$ with PPB approaches

In Figure 6.5, we represent all empirical games for a level of certainty of  $(\eta_v, \eta_b) = (0.5, 0.5)$  where a strategy  $A \in S_{SMS_{EXP3}^\alpha}$  is confronted to a strategy  $B \in S_{PPB}$  (with the exception of EDPE which is strictly dominated by all strategies for every level of certainty). This figure shows us two important points:

- The first point is that deviating from  $B \in S_{PPB}$  to  $A \in S_{SMS_{EXP3}^\alpha}$  in each empirical game is always profitable. Hence, for any strategy  $A \in S_{SMS_{EXP3}^\alpha}$ , the profiles of strategies  $(A, A, A)$  is a Nash equilibrium of the normal-form SAA-inc game in expected utility with set of strategies  $\{A, SB, EPE, EDPE, SCPD\}$  and level of certainty  $(\eta_v, \eta_b) = (0.5, 0.5)$ . Moreover, we observe that deviating from  $B \in S_{PPB}$  to  $A \in$

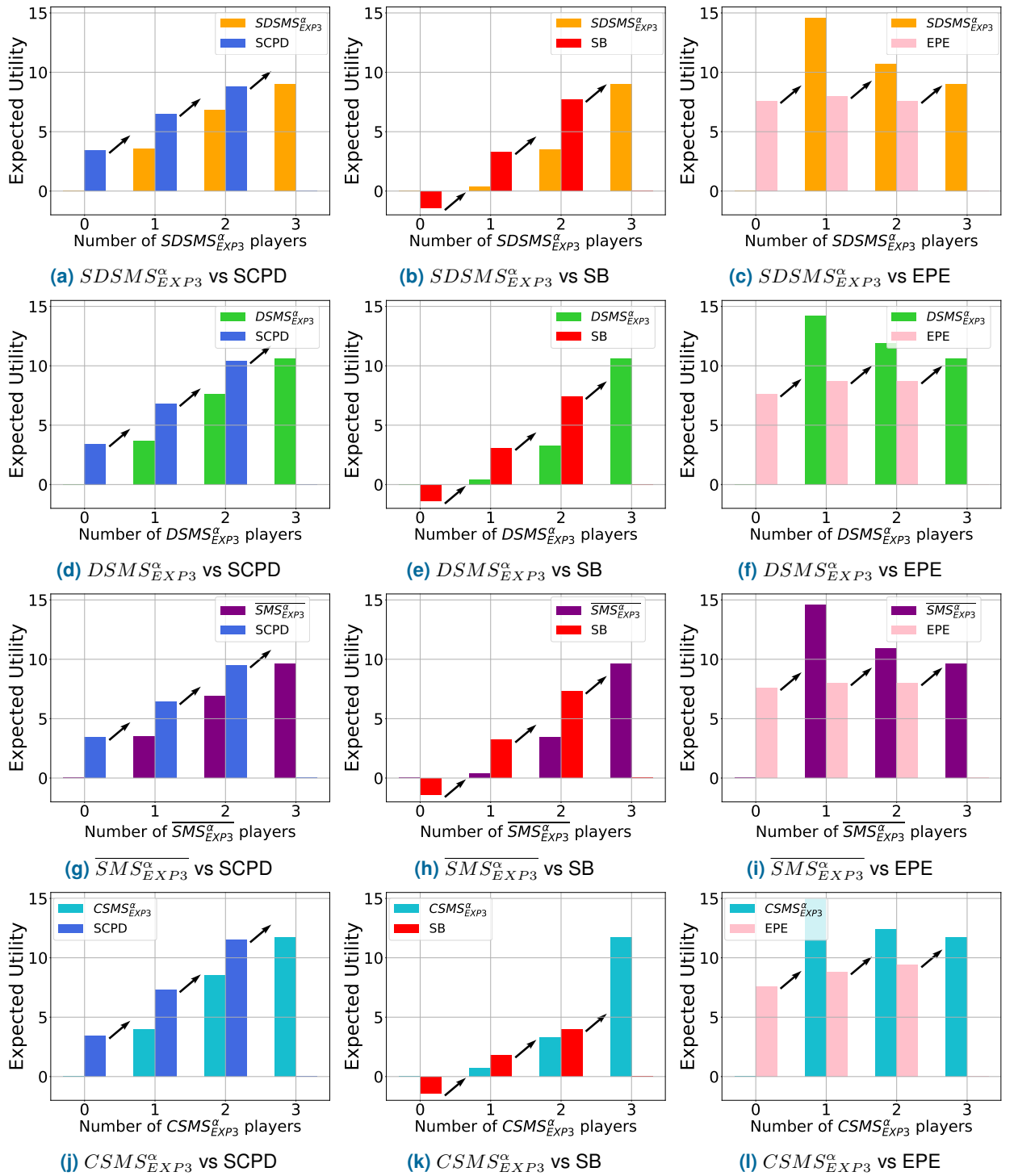
$S_{SMS_{EXP3}^\alpha}$  is also always profitable for the empirical games with level of certainty  $(0, 0)$  and  $(0.8, 0.8)$ . Thus, we can also draw the same conclusions for both of these settings. More generally, it seems that this holds for any level of certainty with  $\eta_v \in [0, 1]$  and  $\eta_v = \eta_b$ .

- The second point is that, although the level of certainty is relatively high, we can notice a lot of similarity between empirical games played by a strategy  $A \in S_{SMS_{EXP3}^\alpha}$  against a same strategy  $B \in S_{PPB}$ . For instance, the empirical games represented in Figure 6.5(a), 6.5(d), 6.5(g) and 6.5(j) are very similar. This highlights the fact that all algorithms in  $S_{SMS_{EXP3}^\alpha}$  heavily rely on  $S_{SMS_{EXP3}^\alpha}$  and, thus, we suspect that similar bidding strategies are computed. This resemblance is even more drastic when the level of certainty is  $(0.8, 0.8)$ . However, when the level of certainty decreases and reaches  $(0, 0)$ , the differences are accentuated, especially with  $CSMS_{EXP3}^\alpha$ .

*Remark.* It is important to note that, for every strategy  $A \in S_{SMS_{EXP3}^\alpha}$ , the hyperparameter  $\alpha$  is chosen such that all deviations from a strategy  $B \in S_{PPB}$  to  $A$  are always profitable. Decreasing  $\alpha$  generally increases the profitability of the deviations and, thus, by taking an  $\alpha$  inferior to the ones proposed in Table 6.2 the same results should be observed. However, by increasing  $\alpha$ , the profitability of some deviations might no longer hold.

Regarding strategy profiles where all bidders play the same strategy, the expected utility when all bidders play the same strategy  $A \in S_{SMS_{EXP3}^\alpha}$  is always greater than the expected utility when all bidders play the same strategy  $B \in S_{PPB}$ . This has been observed for all levels of certainty. For instance, for  $(\eta_v, \eta_b) = (0.5, 0.5)$ ,  $SDSMS_{EXP3}^\alpha$ , which obtains the lowest expected utility amongst strategies in  $S_{SMS_{EXP3}^\alpha}$  when all bidder play the same strategy, has an expected utility 1.17, 1.19 and 2.45 times higher than the expected utility obtained when all bidders play respectively EDPE, EPE and SCPD. If all players decide to play SB, then the resulting expected utility is negative. This highlights the risks undertaken when playing SB as there is fair chance of ending up exposed.

*Remark.* One possible explanation for the relatively good performance of our three determination approaches is that the value distribution of obtaining a set of items  $X$  in our settings progressively resembles a symmetrical unimodal distribution when  $|X|$  increases. For such distributions, the expected value is also the median. Hence,  $\overline{SMS_{EXP3}^\alpha}$  seems to be a relevant approach. Moreover, selecting symmetric values  $\delta$  and  $-\delta$  for generating different combinations of profiles for  $SDSMS_{EXP3}^\alpha$  or  $DSMS_{EXP3}^\alpha$  seems also particularly adapted to this type of distributions. These approaches might be less relevant if we consider



**Fig. 6.5.:** Comparing approaches based on  $SMS_{EXP3}^\alpha$  with PPB approaches through normal-form SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) in expected utility with a level of certainty of  $(\eta_v, \eta_b) = (0.5, 0.5)$ .

bimodal distributions with a significant spread between modes such that the probability of drawing the expected value is zero. This form of distribution could, for example, occur if we believed that an opponent either really desired a set of items  $X$  or not at all (but no in-between). In this specific case, it might be more adapted to select  $\delta$  profiles for each mode separately. Nevertheless, nothing suggests that PPB approaches would be more adapted as they solely rely on the prediction of closing prices which are often computed through expectation methods.

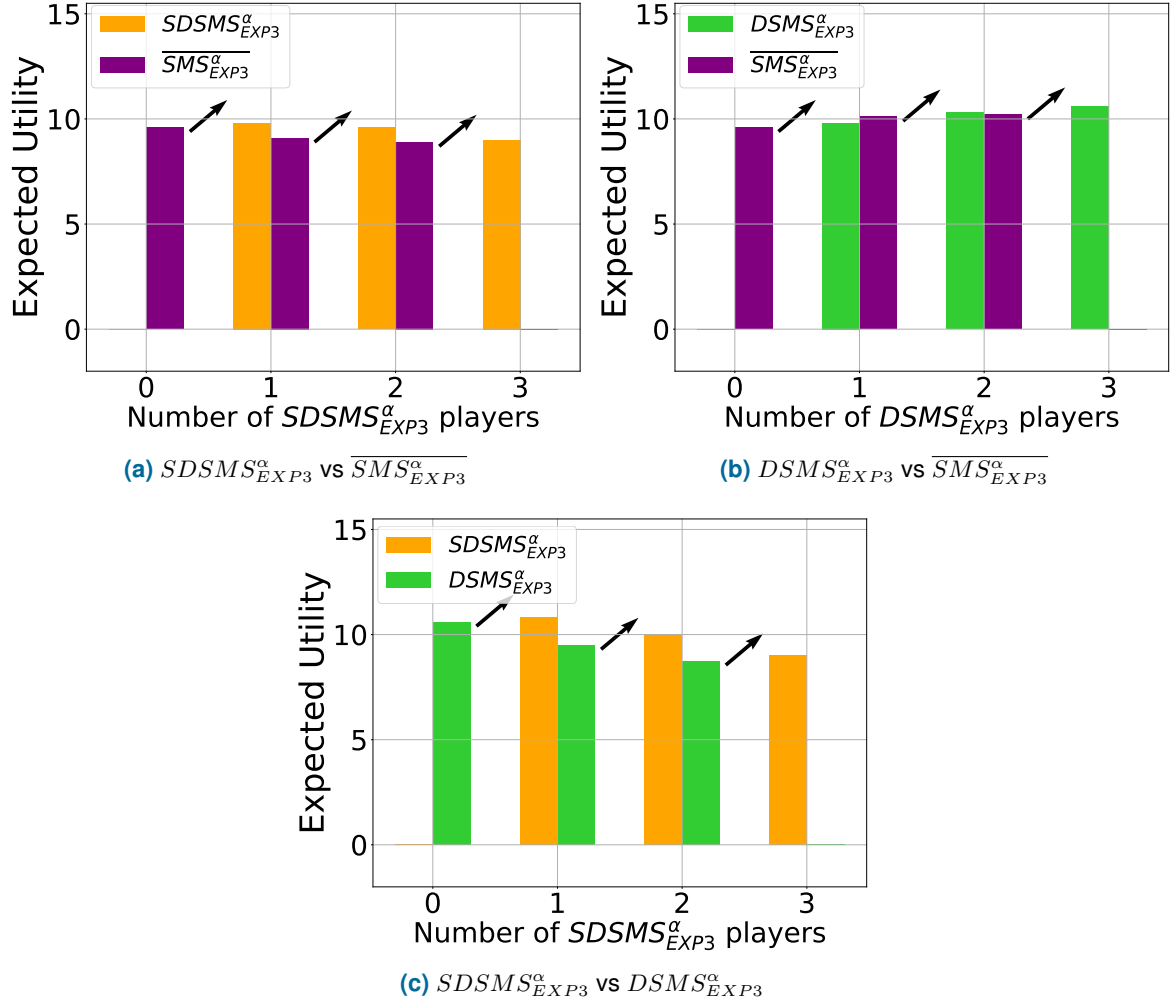
### 6.8.3.2 Comparing our three determinization approaches

In Figure 6.6, we represent all empirical games for a level of certainty of  $(\eta_v, \eta_b) = (0.5, 0.5)$  between our three determinization approaches. We can see that deviating from  $\overline{SMS_{EXP3}^\alpha}$  to either  $SDSMS_{EXP3}^\alpha$  or  $DSMS_{EXP3}^\alpha$  is always profitable. This is also the case for the other two levels of certainty. The fact that these deviations are profitable is not surprising as  $\overline{SMS_{EXP3}^\alpha}$  is our simplest determinization approach which only focuses on the expectation of the opponents' type distribution. Nevertheless, we remain quite surprised that a deviation to the other two determinization approaches never leads to more than a 6% increase in expected utility regardless of the level of certainty. We guess that this could be due to two different factors. First, as mentioned in our last remark, the value distribution of obtaining a set of items  $X$  seems to tend to a symmetrical unimodal distribution when  $|X|$  increases. This may favour methods which only use the expected value of the distribution to make a decision such as  $\overline{SMS_{EXP3}^\alpha}$ . Secondly, the number of different combinations of profiles used in  $SDSMS_{EXP3}^\alpha$  and  $DSMS_{EXP3}^\alpha$  is quite low which might also explain why these algorithms do not significantly outperform  $\overline{SMS_{EXP3}^\alpha}$ . We would like to investigate these two points in future works.

In Figure 6.6(c), we also see that all deviations from  $DSMS_{EXP3}^\alpha$  to  $SDSMS_{EXP3}^\alpha$  are profitable. This is also the case for the two other levels of certainty. Thus, the profile of strategy  $(SDSMS_{EXP3}^\alpha, SDSMS_{EXP3}^\alpha, SDSMS_{EXP3}^\alpha)$  is a Nash equilibrium of the normal-form SAA-inc game in expected utility with strategy set  $\{SDSMS_{EXP3}^\alpha, DSMS_{EXP3}^\alpha, \overline{SMS_{EXP3}^\alpha}, SB, EPE, EDPE, SCPD\}$  for levels of certainty  $(0, 0)$ ,  $(0.5, 0.5)$  and  $(0.8, 0.8)$ . Owing to complementary simulations not shown here, we are relatively confident that this is also the case for all other levels of certainty.

When all bidders play the same strategy,  $DSMS_{EXP3}^\alpha$  always achieves a higher expected utility than  $SDSMS_{EXP3}^\alpha$  or  $\overline{SMS_{EXP3}^\alpha}$  for all levels of certainty. For instance, when  $(\eta_v, \eta_b) = (0.5, 0.5)$ , the expected utility when all players play  $DSMS_{EXP3}^\alpha$  is 1.18 and 1.10 times higher than the expected utility when all players play respectively  $SDSMS_{EXP3}^\alpha$  and

$\overline{SMS}_{EXP3}^\alpha$ . Hence, even though deviating from  $DSMS_{EXP3}^\alpha$  to  $SDSMS_{EXP3}^\alpha$  is always profitable, better coordination between bidders playing the same strategy is obtained with  $DSMS_{EXP3}^\alpha$  than  $SDSMS_{EXP3}^\alpha$ .

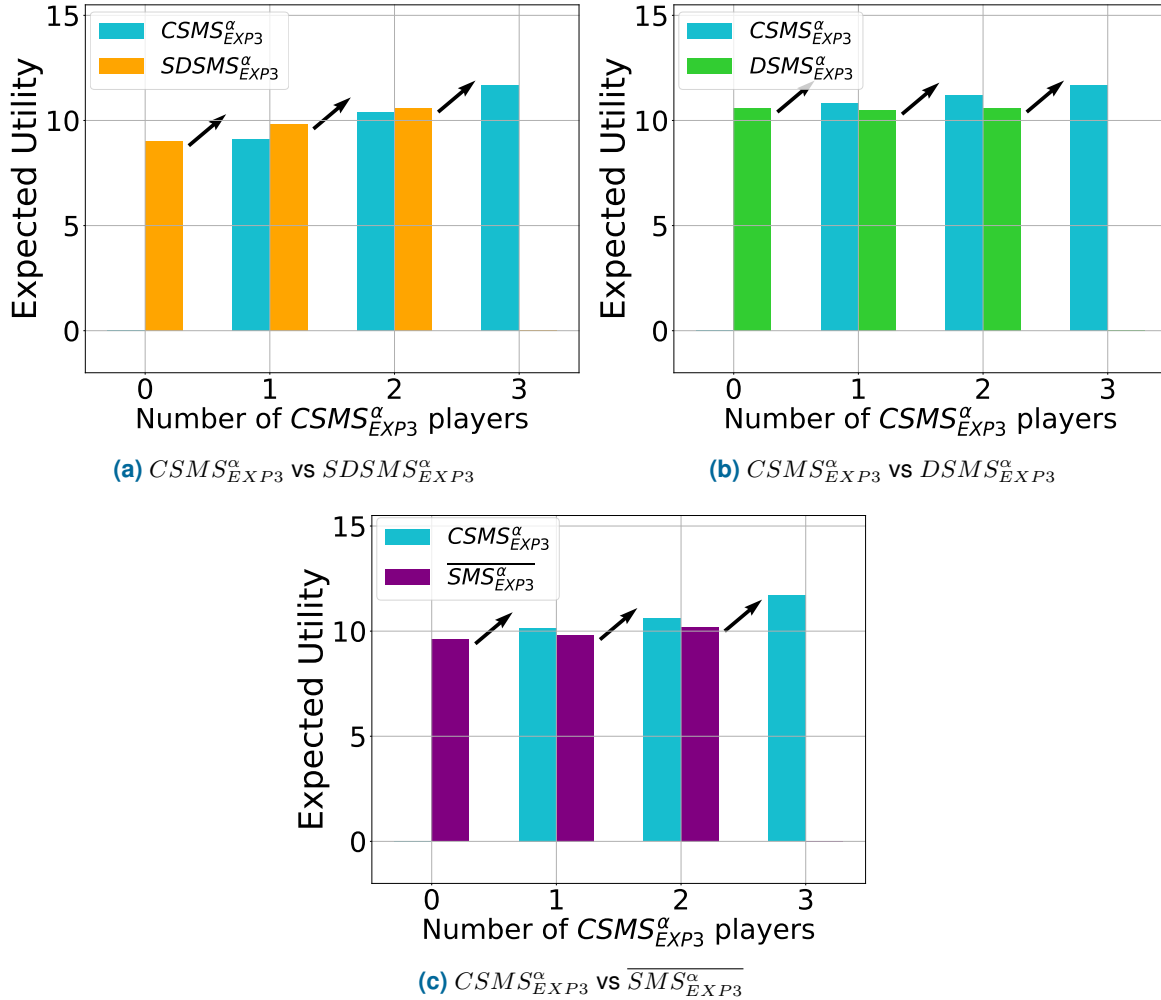


**Fig. 6.6.:** Comparing our three determinization approaches through normal-form SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) in expected utility with a level of certainty of  $(\eta_v, \eta_b) = (0.5, 0.5)$ .

### 6.8.3.3 Impact of uncertainty on our three determinization approaches

In Figure 6.7, we represent the confrontation of each one of our three determinization approaches against  $CSMS_{EXP3}^\alpha$  in a normal-form SAA-inc game in expected utility with  $(\eta_v, \eta_b) = (0.5, 0.5)$ . As expected, in each empirical game, deviating to  $CSMS_{EXP3}^\alpha$  is always profitable. Similar results are observed for the other two levels of certainty. This is due to the fact that a  $CSMS_{EXP3}^\alpha$  bidder knows the exact type of its opponents. Hence, for each level of certainty, the strategy profile  $(CSMS_{EXP3}^\alpha, CSMS_{EXP3}^\alpha, CSMS_{EXP3}^\alpha)$

is a Nash equilibrium of the normal-form SAA-inc game in expected utility with strategy set  $\{CSMS_{EXP3}^\alpha, SDSMS_{EXP3}^\alpha, DSMS_{EXP3}^\alpha, \overline{SMS}_{EXP3}^\alpha, SB, EPE, EDPE, SCPD\}$ . This highlights the fact that having a precise estimate of the opponents' types increases one's performance.



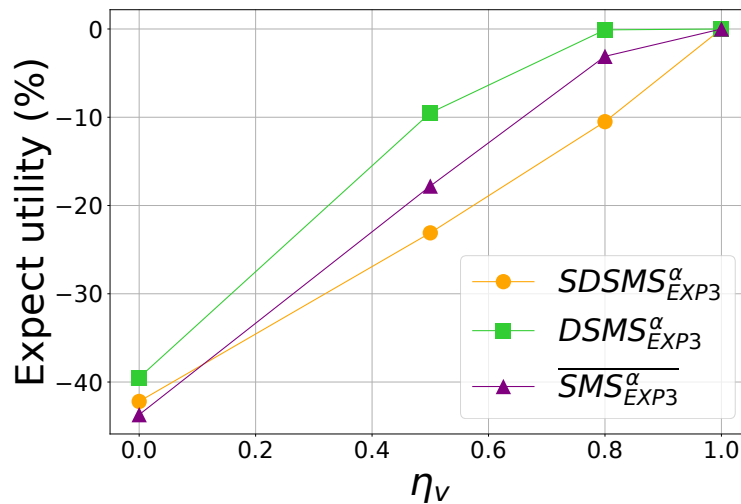
**Fig. 6.7.:** Comparing our three determinization approaches to  $CSMS_{EXP3}^\alpha$  through normal-form SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) in expected utility with a level of certainty of  $(\eta_v, \eta_b) = (0.5, 0.5)$ .

Moreover, uncertainty also highly impacts coordination between bidders playing the same strategy. For instance, we represent in Figure 6.8 the decrease in expected utility (%) when all bidders play the same strategy between  $CSMS_{EXP3}^\alpha$  and a strategy  $A \in \{SDSMS_{EXP3}^\alpha, DSMS_{EXP3}^\alpha, \overline{SMS}_{EXP3}^\alpha\}$  for different levels of certainty. As already mentioned in Section 6.4, when  $\eta_v = \eta_b = 1$ , the bidding game is with complete information and, hence, our three determinization approaches are equivalent to  $CSMS_{EXP3}^\alpha$  (we recall that, for complete information, the set of different combinations of profiles  $L$  is of size 1). In Figure 6.8 we



notice that, the higher the uncertainty is, the more complicated it seems for bidders playing the same determinization approach to coordinate themselves.

Uncertainty impacts our determinization approaches through the outcomes of the rollout phase but also through the initial prediction of closing prices. Whereas  $CSMS_{EXP3}^\alpha$  bidders share the same initial prediction of closing prices, these significantly vary between bidders playing the same strategy  $A$  with  $A \in \{SDSMS_{EXP3}^\alpha, DSMS_{EXP3}^\alpha, \overline{SMS_{EXP3}^\alpha}\}$ , especially for low levels of certainty. This explains the difficulty for bidders playing the same determinization approach to coordinate themselves when uncertainty increases. Moreover, in Figure 6.8, we can see that uncertainty does not impact each determinization approach in the same way. For instance, the decrease in expected utility when all players play  $SDSMS_{EXP3}^\alpha$  compared to when all players play  $CSMS_{EXP3}^\alpha$  seems to be linear with the level of certainty whereas the one corresponding  $DSMS_{EXP3}^\alpha$  seems to be less affected by uncertainty at first but finally ends up at nearly the same point as  $SDSMS_{EXP3}^\alpha$  when  $\eta_v = 0$ .



**Fig. 6.8.:** Comparing the decrease in expected utility (%) obtained when all players play  $CSMS_{EXP3}^\alpha$  and when all players play the same strategy  $A$  with  $A \in \{SDSMS_{EXP3}^\alpha, DSMS_{EXP3}^\alpha, \overline{SMS_{EXP3}^\alpha}\}$  for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with different levels of certainty ( $\eta_v = \eta_b$ ).

#### 6.8.4 Own price effect

Our analysis on own price effect is based on the same performance indicators as in Section 5.5.3.2, i.e. the average price paid per item won and the ratio of items won for each empirical game.

#### 6.8.4.1 Comparing $\overline{SMS_{EXP3}^\alpha}$ with PPB approaches

In this part, we compare  $\overline{SMS_{EXP3}^\alpha}$  to the other four PPB approaches. This is due to the fact that  $\overline{SMS_{EXP3}^\alpha}$  has the highest price paid per item won amongst the four approaches based on  $\overline{SMS_{EXP3}^\alpha}$  against PPB approaches. Hence, the average price paid per item won by  $\overline{SMS_{EXP3}^\alpha}$  against a strategy  $B \in S_{PPB}$  can be seen as an upper-bound of the average price paid per item won by the other two determinization approaches against  $B$ .

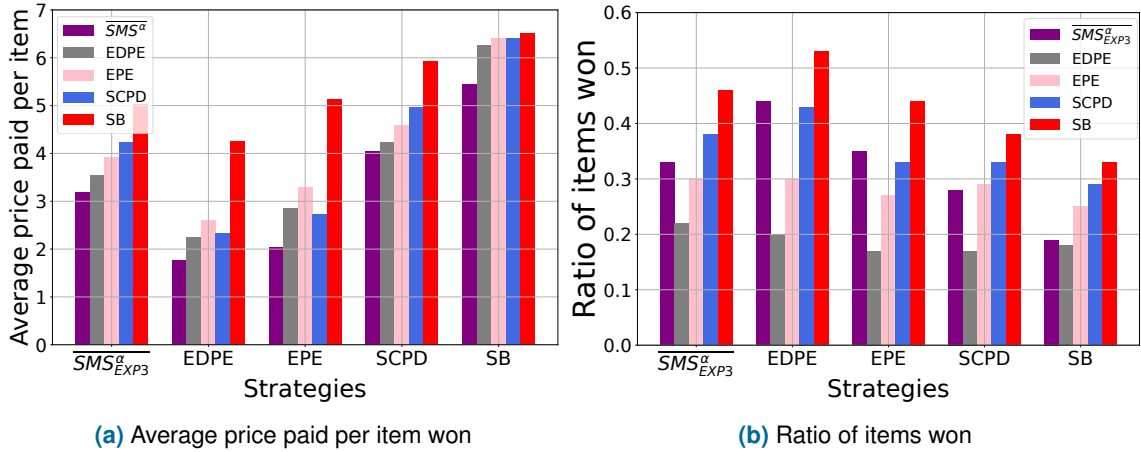
In Figure 6.9(a), we plot the average price paid per item won by each strategy  $A \in S_{PPB} \cup \{\overline{SMS_{EXP3}^\alpha}\}$  against every strategy  $B$  displayed on the x-axis for a level of certainty of (0.5, 0.5). For instance, if  $A = \overline{SMS_{EXP3}^\alpha}$  and  $B = \text{EPE}$ , the average price paid per item won by  $\overline{SMS_{EXP3}^\alpha}$  against EPE is 2.03. It corresponds to the purple bar above index EPE on the x-axis. If  $A = \text{EPE}$  and  $B = \overline{SMS_{EXP3}^\alpha}$ , then the average price paid per item won by EPE against  $\overline{SMS_{EXP3}^\alpha}$  is 3.92. It corresponds to the pink bar above index  $\overline{SMS_{EXP3}^\alpha}$  on the x-axis.

In Figure 6.9(a), we can clearly see that  $\overline{SMS_{EXP3}^\alpha}$  acquires items at a lower price in average than the four PPB approaches against any strategy  $B \in S_{PPB} \cup \{\overline{SMS_{EXP3}^\alpha}\}$ . For instance,  $\overline{SMS_{EXP3}^\alpha}$  spends 29%, 38.4%, 25.3% and 60.4% less per item won against EPE than respectively EDPE, EPE, SCPD and SB. Similar results are also observed for levels of certainty (0, 0) and (0.8, 0.8). Hence, we can conclude that our three determinization approaches pay significantly less per item won than the four PPB approaches regardless of the level of certainty.

To ensure that  $\overline{SMS_{EXP3}^\alpha}$  bidders do not obtain low average prices by only purchasing undesired items, we plot in Figure 6.9(b) the ratio of items won by playing each strategy  $A \in S_{PPB} \cup \{\overline{SMS_{EXP3}^\alpha}\}$  against every strategy  $B$  displayed on the x-axis for a level of certainty of (0.5, 0.5). For instance, if  $A = \overline{SMS_{EXP3}^\alpha}$  and  $B = \text{EDPE}$ , the ratio of items won by  $\overline{SMS_{EXP3}^\alpha}$  against EDPE is 0.44. It corresponds to the purple bar above index EDPE on the x-axis in Figure 6.9(b). If  $A = \text{EDPE}$  and  $B = \overline{SMS_{EXP3}^\alpha}$ , the ratio of items won by EDPE against  $\overline{SMS_{EXP3}^\alpha}$  is 0.22. It corresponds to the grey bar above index  $\overline{SMS_{EXP3}^\alpha}$  on the x-axis in Figure 6.9(b). We can see that each  $\overline{SMS_{EXP3}^\alpha}$  bidder obtains at least 28% of the items against every strategy  $B \in S_{PPB} \cup \{\overline{SMS_{EXP3}^\alpha}\}$  except SB. This shows that  $\overline{SMS_{EXP3}^\alpha}$  computes a competitive strategy.

Regarding strategy profiles where all bidders play the same strategy,  $\overline{SMS_{EXP3}^\alpha}$  has a lower average price paid per item won than EPE, SCPD and SB. Moreover, all items are allocated when all players play  $\overline{SMS_{EXP3}^\alpha}$  whereas only 60.6% of items are allocated when all players

play EDPE. The fact that all items are allocated at a relatively small price explains why the expected utility when all players play  $\overline{SMS}_{EXP3}^\alpha$  is significantly higher than the expected utility when all players play the same PPB strategy. This conclusion can also be drawn for the other two determinization approaches and for the other levels of certainty.



**Fig. 6.9.:** Comparing own price effect of  $\overline{SMS}_{EXP3}^\alpha$  with PPB approaches for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with a level of certainty  $(\eta_v, \eta_b) = (0.5, 0.5)$  through two performance indicators: the average price paid per item won and the ratio of items won

#### 6.8.4.2 Comparing our three determinization approaches

In Figure 6.10, we represent the average price paid per item won and the ratio of items won of each strategy  $A \in S_{SMS_{EXP3}^\alpha} \cup \{SCPD\}$  against every strategy  $B$  displayed on the x-axis for a level of certainty of  $(0.5, 0.5)$ . As the results obtained for our determinization approaches against EPE and SCPD are representative of the performance gap between the three determinization approaches against PPB approaches, we do not represent SB and EDPE on the x-axis. To highlight the performance gap between approaches based on  $SMS_{EXP3}^\alpha$  and PPB approaches (which have already been mentioned in the first part of our analysis on own price effect), we display the results obtained for SCPD (SCPD is used as a representative of PPB approaches).

In order to compare the average price paid per item won by our three determinization approaches displayed in Figure 6.10(a), two different cases need to be considered:

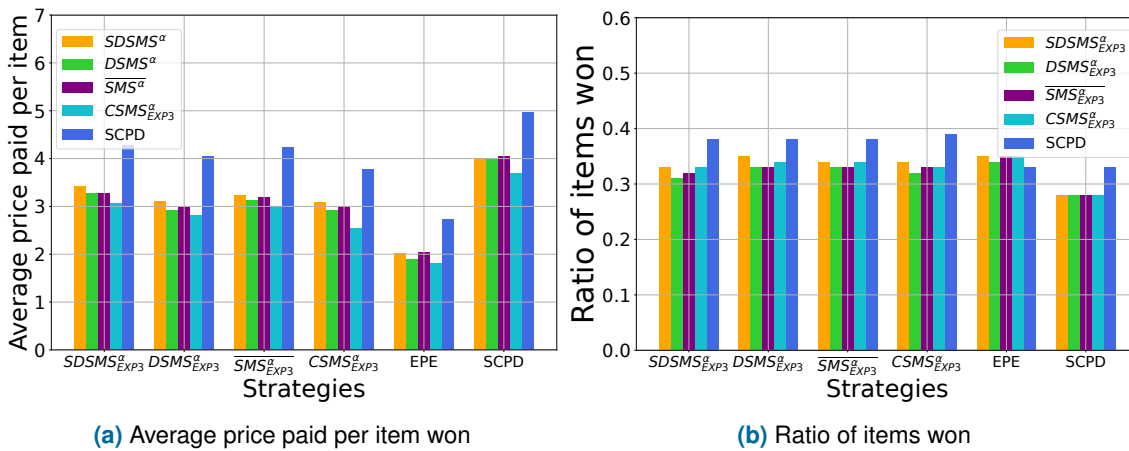
- Against PPB approaches,  $\overline{SMS}_{EXP3}^\alpha$  has the highest average price paid per item won amongst our three determinization approaches.

- Against our three determinization approaches,  $SDSMS_{EXP3}^\alpha$  has the highest average price paid per item won amongst our three determinization approaches.

In both cases, it is always  $DSMS_{EXP3}^\alpha$  which has the lowest average price paid per item won. However, the relative difference between the average price paid per item won never exceeds 7% between two different determinization approaches regardless of the other strategy played.

Regarding the ratio of items won in Figure 6.10(b), it is always  $SDSMS_{EXP3}^\alpha$  which obtains the highest ratio and  $DSMS_{EXP3}^\alpha$  which obtains the lowest ratio. This partially explains why the deviations to  $SDSMS_{EXP3}^\alpha$  from the other determinization approaches are always profitable even though the average price paid per item won is higher.

Regarding strategy profiles where all bidders play the same strategy,  $DSMS_{EXP3}^\alpha$  has an average price paid per item won 8.8% and 15.2% less than respectively  $SMS_{EXP3}^\alpha$  and  $SDSMS_{EXP3}^\alpha$ . Moreover, when all bidders play the same determinization approach, all items are allocated. Similar results are observed for the other levels of certainty. This partially explains why  $DSMS_{EXP3}^\alpha$  has the highest expected utility when all players play the same strategy amongst our three determinization approaches for all levels of certainty.



**Fig. 6.10.:** Comparing own price effect of approaches based on  $SMS_{EXP3}^\alpha$  for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with a level of certainty  $(\eta_v, \eta_b) = (0.5, 0.5)$  through two performance indicators: the average price paid per item won and the ratio of items won

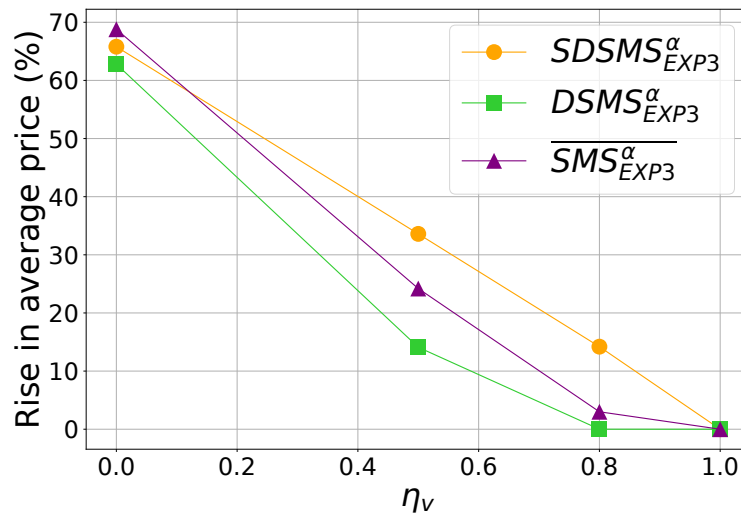
### 6.8.4.3 Impact of uncertainty on our three determinization approaches

In Figure 6.10(a), we can see that  $CSMS_{EXP3}^\alpha$  always pays less in average per item won than our three determinization approaches. For instance,  $CSMS_{EXP3}^\alpha$  spends 10.2%, 6.4%

and 6.4% less against  $SDSMS_{EXP3}^\alpha$  than respectively  $SDSMS_{EXP3}^\alpha$ ,  $DSMS_{EXP3}^\alpha$  and  $\overline{SMS_{EXP3}^\alpha}$ . The fact that  $CSMS_{EXP3}^\alpha$  spends less in average per item won is also the case for the other two levels of certainty. Hence, uncertainty causes a rise in average price paid per item won for our three determinization approaches.

Regarding the ratio of items won in Figure 6.10(b), we do not see any obvious impact of uncertainty on our three determinization approaches.

Uncertainty highly impacts the coordination between bidders playing the same strategy. This is highlighted in Figure 6.11 where we plot the rise in average price paid per item won (%) when all bidders decide to play the same determinization approach instead of  $CSMS_{EXP3}^\alpha$  for different levels of certainty. We see that the average price paid per item won rises when uncertainty increases for our three determinization approaches. Given the fact that for each level of certainty all items are allocated when bidders play the same determinization approach, the rise in average price paid per item won seems to be the principal cause of the decrease in expected utility when uncertainty increases. Indeed, the similarity between Figure 6.11 and Figure 6.8 suggest that the impact of uncertainty on both performance indicators is highly correlated. For instance, in Figure 6.8, the expected utility when all players play  $SDSMS_{EXP3}^\alpha$  increases linearly with the level of certainty while, in Figure 6.11, the average price paid per item won when all players play  $SDSMS_{EXP3}^\alpha$  decreases linearly with the level of certainty.



**Fig. 6.11.:** Comparing the rise in average price paid per item won (%) when all players play  $CSMS_{EXP3}^\alpha$  and when all players play the same strategy  $A$  with  $A \in \{SDSMS_{EXP3}^\alpha, DSMS_{EXP3}^\alpha, \overline{SMS_{EXP3}^\alpha}\}$  for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with different levels of certainty ( $\eta_v = \eta_b$ ).

## 6.8.5 Exposure

As in Section 5.5.3.3, we analyse the exposure problem through two performance indicators: the exposure frequency and the expected exposure.

### 6.8.5.1 Comparing $DSMS_{EXP3}^\alpha$ with PPB approaches

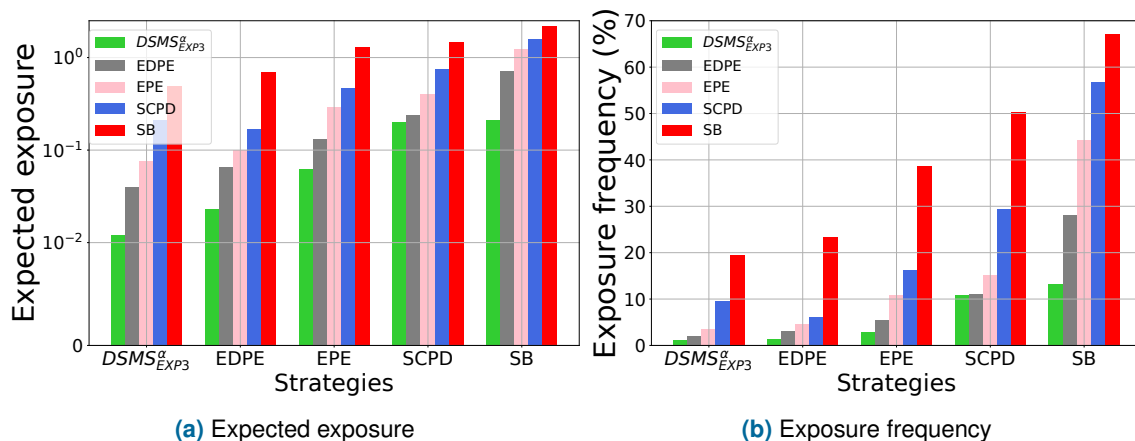
In this part, we compare  $DSMS_{EXP3}^\alpha$  to the other four PPB approaches. This is due to the fact that  $DSMS_{EXP3}^\alpha$  has the highest expected exposure (not necessarily the highest exposure frequency) amongst our determinization approaches against PPB approaches. Hence, the expected exposure suffered by  $DSMS_{EXP3}^\alpha$  against a strategy  $B \in S_{PPB}$  can be seen as an upper-bound of the expected exposure suffered by  $SDSMS_{EXP3}^\alpha$  and by  $\overline{SMS_{EXP3}^\alpha}$  against  $B$ .

We plot in Figure 6.12(a) the expected exposure of each strategy  $A \in S_{PPB} \cup \{DSMS_{EXP3}^\alpha\}$  against every strategy  $B$  displayed on the x-axis. Similarly, we plot in Figure 6.12(b) the exposure frequency of each strategy  $A \in S_{PPB} \cup \{DSMS_{EXP3}^\alpha\}$  against every strategy  $B$  displayed on the x-axis. For instance, if  $A = DSMS_{EXP3}^\alpha$  and  $B = SB$ , the expected exposure and exposure frequency are respectively 0.21 and 13.3%. They both correspond respectively to the green bar above SB index on the x-axis in Figure 6.12(a) and in Figure 6.12(b). All these results are obtained for a level of certainty of  $(\eta_v, \eta_b) = (0.5, 0.5)$

We can see in Figure 6.12 that  $DSMS_{EXP3}^\alpha$  has the lowest expected exposure and exposure frequency against every strategy displayed on the x-axis. For instance,  $DSMS_{EXP3}^\alpha$  has 2.1, 4.7, 7.6 and 20.8 times less expected exposure against EPE than respectively EDPE, EPE, SCPD and SB. Moreover,  $DSMS_{EXP3}^\alpha$  has 2, 3.9, 5.8 and 13.8 less chance of ending up exposed against EPE than respectively EDPE, EPE, SCPD and SB. Hence, we can conclude that our determinization approaches suffer less from exposure than PPB approaches.

Regarding strategy profiles when all bidders play the same strategy, the one corresponding to  $DSMS_{EXP3}^\alpha$  induces the lowest expected exposure and exposure frequency. For instance, when all bidders play  $DSMS_{EXP3}^\alpha$  there is 1.2% chance that a bidder suffers from exposure whereas, if they all play EPE, there is 10.9% chance. Moreover, in Section 5.5.3.3, we had seen that, in the case of complete information, when all bidders play EPE there is 0.1% chance that a bidder suffers from exposure. This was partially due to the fact that EPE bidders shared the same initial prediction of closing price if the induced bidding game is

with complete information. However, in the case of incomplete information, this is no longer the case. Hence, the expected exposure and exposure frequency when all bidders play EPE increase drastically. This highlights the fact that solely relying on an initial prediction of closing prices such as EPE (or other PPB approaches) is risky, especially in the case of high uncertainty.



**Fig. 6.12.:** Comparing the exposure suffered by  $DSMS_{EXP3}^{\alpha}$  with the exposure suffered by PPB approaches for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with a level of certainty  $(\eta_v, \eta_b) = (0.5, 0.5)$  through two performance indicators: the expected exposure and the exposure frequency

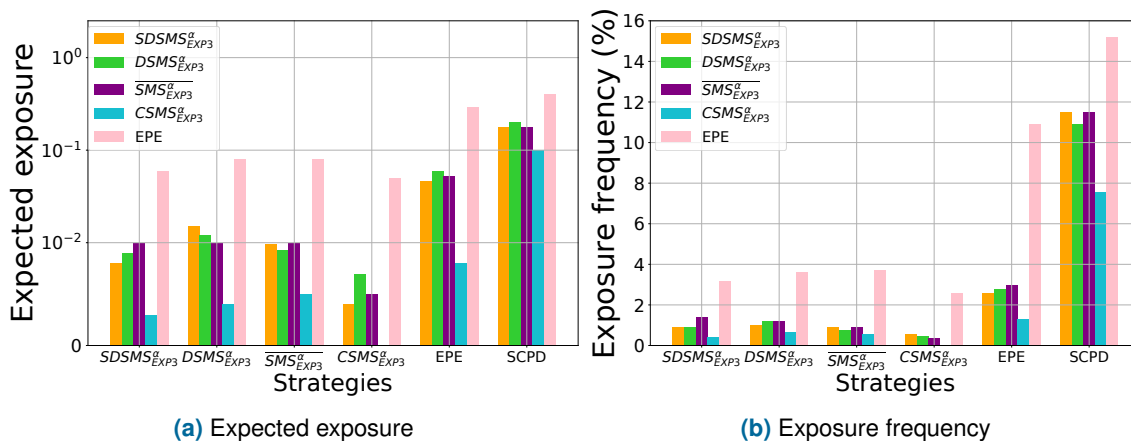
### 6.8.5.2 Comparing our three determinization approaches

In Figure 6.13, we represent the expected exposure and the exposure frequency of each strategy  $A \in S_{SMS_{EXP3}^{\alpha}} \cup \{EPE\}$  against every strategy  $B$  displayed on the x-axis for a level of certainty of  $(0.5, 0.5)$ . As the results obtained for our determinization approaches against EPE and SCPD are representative of the performance gap between the three determinization approaches against PPB approaches, we do not represent SB and EDPE on the x-axis. To highlight the performance gap regarding exposure between approaches based on  $SMS_{EXP3}^{\alpha}$  and PPB approaches, we also display the results for EPE.

Although  $DSMS_{EXP3}^{\alpha}$  is the strategy amongst our determinization approaches with the lowest average price paid per item won, it is also the most risky strategy amongst our determinization approaches against PPB approaches as it incurs the highest expected exposure. However, it does not necessarily obtains the highest exposure frequency as we can see against EPE and SCPD.

Nothing particular can be said against approaches based on  $SMS_{EXP3}^{\alpha}$  as no systematic ordering between our determinization approaches can be observed.

When all bidders play the same strategy,  $SDSMS_{EXP3}^\alpha$  obtains the lowest expected exposure and exposure frequency amongst our three determinization approaches. It is important to note that the highest exposure frequency is only of 1.2% and obtained when all bidders play  $DSMS_{EXP3}^\alpha$ .



**Fig. 6.13.:** Comparing the expected exposure and exposure frequency of our approaches based on  $SMS_{EXP3}^\alpha$  for SAA-inc games ( $n = 3, m = 9, \varepsilon = 1$ ) with a level of certainty  $(\eta_v, \eta_b) = (0.5, 0.5)$

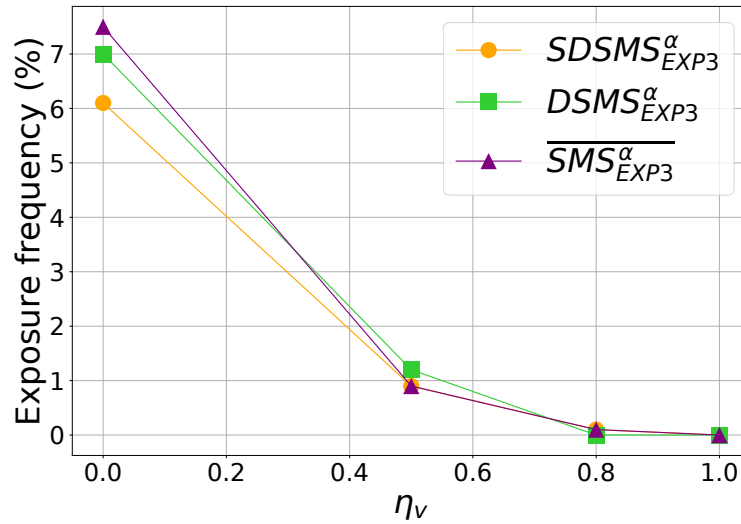
### 6.8.5.3 Impact of uncertainty on our three determinization approaches

In Figure 6.13, we can see that  $CSMS_{EXP3}^\alpha$  suffers far less exposure against any strategy  $B$  on the x-axis than our three determinization approaches. For instance,  $CSMS_{EXP3}^\alpha$  incurs 5.8, 7.5 and 6.6 times less expected exposure and has 50%, 54% and 57% less chance of ending up exposed against EPE than respectively  $SDSMS_{EXP3}^\alpha$ ,  $DSMS_{EXP3}^\alpha$  and  $SMS_{EXP3}^\alpha$ . Moreover, when all bidders play the same strategy,  $CSMS_{EXP3}^\alpha$  never leads to exposure. Hence, uncertainty rises the risk of exposure for our three determinization approaches.

In Figure 6.14, we plot the exposure frequency of our three determinization approaches when all bidders play the same strategy for different levels of certainty. We remark that our three determinization approaches are impacted by uncertainty in nearly the same way in the sense that the exposure frequency drastically drops from  $\eta_v = 0$  to  $\eta_v = 0.5$  and then slowly decreases to 0 between  $\eta_v = 0.5$  and  $\eta_v = 1$ . Hence, the coordination between bidders playing the same strategy becomes more risky when uncertainty increases. Moreover, it is interesting to highlight that, even though  $DSMS_{EXP3}^\alpha$  obtains higher expected utility than  $SDSMS_{EXP3}^\alpha$  when all bidders play the same strategy as bidders acquire items at a lower



price in average, it also leads to more exposure.



**Fig. 6.14.:** Exposure frequency (%) of our three determinization approaches when all players play the same strategy for SAA-inc games ( $n = 3$ ,  $m = 9$ ,  $\varepsilon = 1$ ) with different levels of certainty ( $\eta_v = \eta_b$ ).

### 6.8.6 Lessons learned from the experiments

Through these extensive experiments, we have seen that:

- Our three determinization approaches significantly outperform PPB approaches for all levels of certainty, notably by better tackling the own price effect and the exposure problem in budget and eligibility constrained environments.
- Differences between our three determinization approaches are less notable than with PPB approaches. We observe that  $SDSMS_{EXP3}^{\alpha}$  slightly outperforms  $DSMS_{EXP3}^{\alpha}$  and  $\overline{SMS}_{EXP3}^{\alpha}$  for all levels of certainty. However, if all players decide to play the same strategy, then playing  $DSMS_{EXP3}^{\alpha}$  achieves higher expected utility than  $SDSMS_{EXP3}^{\alpha}$ .
- Uncertainty impacts our three determinization approaches by increasing their average price paid per item won as well as their risk of exposure. This causes a decrease in their respective expected utility. Moreover, when all bidders are playing the same strategy, uncertainty highly impacts their coordination, especially in the case of wide type distributions.

## 6.9 Conclusion

In this chapter, we introduce three determinization approaches based on  $SMSS_{EXP3}^\alpha$  that tackles simultaneously the *exposure problem*, the *own price effect*, *budget constraints*, the *eligibility management problem* for a simplified version of SAA (SAA-inc) with incomplete information. Each approach updates its belief about opponents' budgets by tracking bid exposure. Experiments are run on instances generated by our new framework which extends the experimental settings of the last two chapters to incomplete information. On these instances, we show that each determinization approach largely outperforms state-of-the-art algorithms. Even though smaller instances are considered in this chapter compared to the last chapter, our determinization approaches can easily be applied to instances of realistic size as they essentially rely on  $SMSS_{EXP3}^\alpha$ .

Amongst our three determinization approaches,  $SDSMSS_{EXP3}^\alpha$  seems slightly better than the two others by achieving higher expected utility. However, better coordination between bidders is obtained for  $DSMSS_{EXP3}^\alpha$ . Moreover, only considering the expected value of each type distribution and applying  $SMSS_{EXP3}^\alpha$  to the resulting determinized SAA-inc game gives surprisingly good results which are comparable with the other two determinization approaches. Although we think that these determinization approaches can easily be improved by increasing the number of generated profiles or by modifying their selection phase with an even better selection strategy, we believe that the greatest area for improvement lies in inference methods in order to obtain narrower type distributions.

In future works, we would also like to investigate why  $SDSMSS_{EXP3}^\alpha$  has weaker coordination than  $DSMSS_{EXP3}^\alpha$  when all bidders play the same strategy. In the short term, our aim would be to achieve at least the same coordination as  $DSMSS_{EXP3}^\alpha$  while conserving the other results obtained for  $SDSMSS_{EXP3}^\alpha$ .



# Conclusion

## Contents

---

7.1	Summary of achievements . . . . .	199
7.2	Short-term prospects . . . . .	202
7.3	Long-term prospects . . . . .	203

---

## 7.1 Summary of achievements

In this thesis, we have focused on adversarial search methods to solve the bidding problem in SAA. By pursuing this research direction, we were able to propose an efficient bidding algorithm for the original format of SAA with incomplete information that tackles simultaneously its four main strategical issues: the *exposure problem*, the *own price effect*, *budget constraints* and the *eligibility management problem*. Our solution is based on an MCTS which relies on a new method of prediction of closing prices. Through the use of scalarised rewards, our algorithm allows a bidder to arbitrate between expected-utility and risk-aversion. Moreover, through typical examples taken from the literature and extensive numerical experiments on instances of realistic size, we have shown that our solution significantly outperforms state-of-the-art bidding strategies. Unlike preceding related works, no specific assumption is made on the form of the value functions such as super-additivity, making our approach more generic.

To achieve such results, we have proceeded step-by-step by increasing progressively the complexity of the induced bidding game. Throughout this thesis, we have introduced three different simplified SAA models: d-SAA, SAA-c and SAA-inc. We summarize hereafter our various contributions to the existing literature for each one of these models.

- In d-SAA,
  - We compute the information set space complexity as well as the game tree complexity of the induced bidding game.

- We introduce a new concept for the prediction of final auction prices named *frontier prediction of final prices*. Unlike preceding prediction methods, it meets the three following desirable properties: (1) it exists in all auctions, (2) it takes into account auctions' particularities and (3) the final prediction is independent of single specific strategy profile.
  - We present a simple method based on a specific sequence which converges to the *frontier prediction of final prices*. We prove its convergence in simple environments (Property 4.3) and we conjecture this result in more complex environments (Conjecture 4.1).
  - We present the first efficient bidding algorithm  $MS^\lambda$  that tackles simultaneously the own price effect and the exposure problem.
  - We propose a general framework to generate value functions without making any specific assumption on its form.
  - We show that  $MS^\lambda$  significantly outperforms state-of-the-art bidding strategies on instances with two players.
- In SAA-c,
    - We compute the information set space complexity as well as the game tree complexity of the induced bidding game with no activity rules and unlimited budgets.
    - We extend PP strategy to budget and eligibility constrained environments.
    - We adapt our preceding prediction method for final auction prices based on a specific sequence to budget constraints, eligibility constraints and stochasticity. We prove its convergence in *Example 1* (see Section 2.2.3) and conjecture this result in the general case (Conjecture 5.1).
    - We present the first efficient bidding strategy  $SMS^\alpha$  that tackles simultaneously the four main strategical issues of SAA. Moreover, we introduce a hyperparameter  $\alpha$  that allows a bidder to arbitrate between expected-utility and risk-aversion.
    - We extend our preceding general framework to the generation of budgets.

- Through typical examples taken from the literature and extensive numerical experiments on instances of realistic size, we show that  $SMS^\alpha$  outperforms state-of-the-art bidding strategies by better tackling the exposure problem and the own price effect in budget and eligibility constrained environments.
- We provide an analysis of the impact of  $\alpha$  on  $SMS^\alpha$ .
- In SAA-inc,
  - We compute the information set space complexity as well as the game tree complexity of the induced bidding game with no activity rules and unlimited budgets.
  - We improve our preceding algorithm  $SMS^\alpha$  by replacing the UCT selection phase with the EXP3 algorithm.
  - We propose three different determinization approaches based on  $SMS_{EXP3}^\alpha$ :  $\overline{SMS}_{EXP3}^\alpha$ ,  $DSMS_{EXP3}^\alpha$  and  $SDSMS_{EXP3}^\alpha$ . These are the first efficient bidding algorithms that tackle simultaneously the four main strategical issues of SAA in an environment with incomplete information.
  - We propose a simple inference method based on tracking bid exposure for updating one's belief about an opponent's budget.
  - We propose a general experimental framework for generating types with a certain level of certainty that extends our preceding framework to incomplete information.
  - Through extensive experiments, we show that our three determinization approaches significantly outperform state-of-the-art bidding strategies by better tackling the own price effect and the exposure problem in budget and eligibility constrained environments with incomplete information. Moreover, we notice that, by playing  $SDSMS_{EXP3}^\alpha$ , a player obtains a slightly higher expected utility than with the other two determinization approaches. However, better coordination between bidders is obtained with  $DSMS_{EXP3}^\alpha$ . Furthermore, we show that uncertainty impacts our three determinization approaches by notably reducing their expected utility.

## 7.2 Short-term prospects

With a little more time, we would have liked to investigate some results obtained in SAA-inc, including examining the factors that lead  $DSMS_{EXP3}^\alpha$  to achieve a higher expected utility than  $SDSMS_{EXP3}^\alpha$  when all bidders play the same strategy. Furthermore, we had intended to analyse the impact of increasing the number of different combinations of profiles for both of these determinization approaches. For example, to what extent the performance can be improved by adding more profiles to the set of different combinations of profiles  $L$ . Furthermore, although we are pretty confident that the results would be similar to the ones already obtained, we would have preferred to apply our determinization approaches to larger instances.

Another perspective for future works would be to improve our algorithms by modifying some of its MCTS search iteration phases. For instance, other selection strategies than UCT or the EXP3 algorithm can be considered such as *Regret Matching* [Tak, 2014].

Concerning our method for predicting final auction prices, as it is relatively time-consuming and prevents us from considering many combinations of types, it could be interesting to see if a relation can be found between types and our final prediction. For instance, to what extent our prediction of final prices is impacted when a player's value function or budget is slightly modified. Moreover, anticipating how our prediction of final prices evolves when, for example, a player's budget increases or decreases may help one to build a meaningful interpolation model. If such an interpolation model can be developed, we could compute an initial prediction of final prices for a few representative types that span the entire distribution of types, and then use our interpolation model to estimate final prices for all types. From there on, instead of having a small set of combinations of profiles,  $SDSMS_{EXP3}^\alpha$  could sample randomly from the whole types' distribution at each search iteration.

Although we have obtained very good results by combining MCTS with our method for predicting final prices, one of the biggest motivations for using this method has yet to be proven. Indeed, we have only proven the convergence of our method on small instances and not in the general case. This remains an open issue. A proof direction that we find promising is provided below Conjecture 4.1.

More complex SAA models can also be considered with additional constraints. For example, one could model spectral caps in SAA-inc by imposing a maximal number of items a bidder

can get. One big advantage with our determinization approaches is that integrating such constraints is fairly easy. Moreover, another relevant extension of our SAA-inc model would be to integrate the items won and the price paid by other bidders in a player's utility. By doing so, one can model situations in which a player prefers one bidder to obtain a specific item over another at the same price. Furthermore, in this thesis, a bidder never has any interest in decreasing their opponents' utility. In this extension, however, driving up the prices of items won by opponents increases a bidder's utility. Hence, a new strategic behaviour emerges.

### 7.3 Long-term prospects

Although we have already discussed some potential enhancements of our algorithms in the short-term prospects, we firmly believe that the most significant area of improvement concerns inference methods. Indeed, as seen in our extensive experimental results, uncertainty highly impacts our algorithms as well as PPB approaches. Hence, we believe that narrowing the type distributions would potentially lead to a greater increase in performance than just improving one of our MCTS search iteration phases. However, inferring one's value functions by just observing its bids seems a difficult or even a nearly impossible task in complex environments, especially if we consider the possibility that one is allowed to bluff. Hence, one may need to consider a set of hypothesis on one's bidding behaviour as well as some exogenous variables in order to have a chance to narrow its type distribution. Moreover, with the assumption that players bid sincerely, we think that a good starting point for an inference method on value functions is to use the inequality given by the *revealed preference activity rule* in Definition 1.1.

One interesting approach for future work would be to collect data on real passed SAA and try to replay the auctions with our algorithms. Since our algorithms take either point-wise estimations of types or type distributions as inputs, a pressing question arises about how to generate these inputs from the gathered data. Moreover, if such inputs are somehow created, one could try to analyse which strategy the real bidders follow. For instance, do real bidders play generally straightforwardly? Would there be an increase in their utility if they had bid according to the strategy computed by one of our algorithms? If answers are found, this could be quite valuable information for mobile operators regarding future auctions.



Finally, one could extend our algorithms to other multi-round auction mechanism. For instance, an intriguing problem would be to consider multiple SAA that are held sequentially where bidders have a fixed budget for all auctions. This could model real-life situations where groups of licenses are auctioned successively based on their bandwidth, involving the same set of bidders in each auction. Hence, bidders need to plan how to allocate their budget in each different auction. Indeed, spending all its budget in the first auction or saving up all its budget for the last auction might not be strategically optimal. Moreover, these decisions can be made dynamically at each round of an auction depending on the actual bid prices of the considered licences and the budget already spent by the other bidders.

# Bibliography

- [Abramson, 1990] Bruce Abramson. “Expected-outcome: A general model of static evaluation”. In: *IEEE transactions on pattern analysis and machine intelligence* 12.2 (1990), pp. 182–193 (cit. on p. 76).
- [Allis, 1994] Louis Victor Allis et al. *Searching for solutions in games and artificial intelligence*. Ponsen & Looijen Wageningen, 1994 (cit. on pp. 38–40).
- [ANACOM, 2021] ANACOM. *ANACOM commences procedure to amend the 5G auction regulation in order to prohibit use of smaller bidding increments*. <https://www.anacom.pt/render.jsp?contentId=1695741>. ANACOM. 2021 (cit. on pp. 16, 22).
- [Arcep, 2020] Arcep. *5G*. <https://www.arcep.fr/actualites/actualites-et-communiques/detail/n/5g-011020.html>. Arcep. 2020 (cit. on p. 9).
- [Arcep, 2021] Arcep. *Fréquences-Outre-mer*. <https://www.arcep.fr/actualites/actualites-et-communiques/detail/n/frequences-outr-mer-030821.html>. Arcep. 2021 (cit. on p. 20).
- [Arneson, 2010] Broderick Arneson, Ryan B Hayward, and Philip Henderson. “Monte Carlo tree search in Hex”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 2.4 (2010), pp. 251–258 (cit. on pp. 82, 83).
- [Arrow, 1971] Kenneth Joseph Arrow, Frank Hahn, et al. “General competitive analysis”. In: (1971) (cit. on pp. 51, 101).
- [Auer, 1995] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. “Gambling in a rigged casino: The adversarial multi-armed bandit problem”. In: *Proceedings of IEEE 36th annual foundations of computer science*. IEEE. 1995, pp. 322–331 (cit. on pp. 78, 80).

- [Auer, 2002a] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. “Finite-time analysis of the multiarmed bandit problem”. In: *Machine learning* 47.2 (2002), pp. 235–256 (cit. on pp. 78, 80).
- [Auer, 2002b] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. “The nonstochastic multiarmed bandit problem”. In: *SIAM journal on computing* 32.1 (2002), pp. 48–77 (cit. on pp. 78, 159).
- [Ausubel, 2006a] Lawrence M Ausubel, Peter Cramton, and Paul Milgrom. “The clock-proxy auction: A practical combinatorial auction design”. In: *Handbook of Spectrum Auction Design* (2006), pp. 120–140 (cit. on pp. 25, 27, 28).
- [Ausubel, 2006b] Lawrence M Ausubel, Paul Milgrom, et al. “The lovely but lonely Vickrey auction”. In: *Combinatorial auctions* 17 (2006), pp. 22–26 (cit. on p. 26).
- [Ausubel, 2011] Lawrence M Ausubel, Peter Cramton, et al. “Activity rules for the combinatorial clock auction”. In: *University of Maryland, Department of Economics, Maryland* (2011) (cit. on p. 27).
- [Ausubel, 2014] Lawrence M Ausubel, Peter Cramton, Marek Pycia, Marzena Rostek, and Marek Weretka. “Demand reduction and inefficiency in multi-unit auctions”. In: *The Review of Economic Studies* 81.4 (2014), pp. 1366–1400 (cit. on p. 42).
- [Ausubel, 2017] Lawrence M Ausubel and Oleg Baranov. *A practical guide to the combinatorial clock auction*. 2017 (cit. on pp. 25, 27–29).
- [Baier, 2014] Hendrik Baier and Mark HM Winands. “MCTS-minimax hybrids”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 7.2 (2014), pp. 167–179 (cit. on p. 82).
- [Barrett, 2008] Leon Barrett and Srinu Narayanan. “Learning all optimal policies with multiple criteria”. In: *Proceedings of the 25th international conference on Machine learning*. 2008, pp. 41–47 (cit. on p. 132).
- [Bazerman, 1983] Max H Bazerman and William F Samuelson. “I won the auction but don’t want the prize”. In: *Journal of conflict resolution* 27.4 (1983), pp. 618–634 (cit. on p. 14).
- [Bichler, 2017] Martin Bichler and Jacob K Goeree. “Frontiers in spectrum auction design”. In: *International Journal of Industrial Organization* 50 (2017), pp. 372–391 (cit. on pp. 36, 43).
- [Bikhchandani, 1997] Sushil Bikhchandani and John W Mamer. “Competitive equilibrium in an exchange economy with indivisibilities”. In: *Journal of economic theory* 74.2 (1997), pp. 385–413 (cit. on p. 48).

- [Binmore, 2002] Ken Binmore and Paul Klemperer. “The biggest auction ever: the sale of the British 3G telecom licences”. In: *The Economic Journal* 112.478 (2002), pp. C74–C96 (cit. on p. 48).
- [Blumrosen, 2007] Liad Blumrosen and Noam Nisan. “Combinatorial auctions”. In: *Algorithmic game theory* 267 (2007), p. 300 (cit. on p. 50).
- [Blumrosen, 2023] Liad Blumrosen and Eilon Solan. “Selling spectrum in the presence of shared networks: The case of the Israeli 5G auction”. In: *Telecommunications Policy* 47.2 (2023), p. 102481 (cit. on p. 30).
- [Bouzy, 2004] Bruno Bouzy and Bernard Helmstetter. “Monte-carlo go developments”. In: *Advances in computer games*. Springer, 2004, pp. 159–174 (cit. on p. 76).
- [Browne, 2012] Cameron B Browne, Edward Powley, Daniel Whitehouse, Simon M Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. “A survey of monte carlo tree search methods”. In: *IEEE Transactions on Computational Intelligence and AI in games* 4.1 (2012), pp. 1–43 (cit. on pp. 10, 39, 81, 83).
- [Brusco, 2002] Sandro Brusco and Giuseppe Lopomo. “Collusion via signalling in simultaneous ascending bid auctions with heterogeneous objects, with and without complementarities”. In: *The Review of Economic Studies* 69.2 (2002), pp. 407–436 (cit. on pp. 41, 57, 114, 115).
- [Brusco, 2009] Sandro Brusco and Giuseppe Lopomo. “Simultaneous ascending auctions with complementarities and known budget constraints”. In: *Economic Theory* 38.1 (2009), pp. 105–124 (cit. on pp. 58, 115).
- [Bulow, 2009] Jeremy Bulow, Jonathan Levin, and Paul Milgrom. *Winning play in spectrum auctions*. Tech. rep. National Bureau of Economic Research, 2009 (cit. on pp. 47, 48, 59, 173).
- [Bundesnetzagentur, 2022] Bundesnetzagentur. *Spectrum for wireless access for the provision of telecommunications services*. [https://www.bundesnetzagentur.de/EN/Areas/Telecommunications/Companies/FrequencyManagement/ElectronicCommunicationsServices/ElectronicCommunicationServices\\_node.html](https://www.bundesnetzagentur.de/EN/Areas/Telecommunications/Companies/FrequencyManagement/ElectronicCommunicationsServices/ElectronicCommunicationServices_node.html). Bundesnetzagentur. 2022 (cit. on pp. 22, 31).
- [Bünz, 2022] Benedikt Bünz, Benjamin Lubin, and Sven Seuken. “Designing core-selecting payment rules: A computational search approach”. In: *Information Systems Research* 33.4 (2022), pp. 1157–1173 (cit. on p. 26).

- [Bykowsky, 2000] Mark M Bykowsky, Robert J Cull, and John O Ledyard. “Mutually destructive bidding: The FCC auction design problem”. In: *Journal of Regulatory Economics* 17.3 (2000), pp. 205–228 (cit. on pp. 15, 56).
- [Campbell, 2002] Murray Campbell, A Joseph Hoane Jr, and Feng-hsiung Hsu. “Deep blue”. In: *Artificial intelligence* 134.1-2 (2002), pp. 57–83 (cit. on p. 75).
- [Chaslot, 2008a] Guillaume M Jb Chaslot, Mark HM Winands, H Jaap van den Herik, Jos WHM Uiterwijk, and Bruno Bouzy. “Progressive strategies for Monte-Carlo tree search”. In: *New Mathematics and Natural Computation* 4.03 (2008), pp. 343–357 (cit. on pp. 81, 82).
- [Chaslot, 2008b] Guillaume MJ -B Chaslot, Mark HM Winands, and H Jaap van Den Herik. “Parallel monte-carlo tree search”. In: *International conference on computers and games*. Springer. 2008, pp. 60–71 (cit. on p. 86).
- [Cheng, 2003] Shih-Fen Cheng, Evan Leung, Kevin M Lochner, Kevin O’Malley, Daniel M Reeves, L Julian Schwartzman, and Michael P Wellman. “Walverine: A Walrasian trading agent”. In: *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*. 2003, pp. 465–472 (cit. on p. 51).
- [Childs, 2008] Benjamin E Childs, James H Brodeur, and Levente Kocsis. “Transpositions and move groups in Monte Carlo tree search”. In: *IEEE Symposium On Computational Intelligence and Games*. IEEE. 2008, pp. 389–395 (cit. on pp. 88, 136).
- [Chowdhury, 2018] Moinul Morshed Porag Chowdhury, Christopher Kiekintveld, Son Tran, and William Yeoh. “Bidding in Periodic Double Auctions Using Heuristics and Dynamic Monte Carlo Tree Search.” In: *IJCAI*. 2018, pp. 166–172 (cit. on p. 88).
- [Coase, 1959] Ronald H Coase. “The federal communications commission”. In: *The Journal of Law and Economics* 2 (1959), pp. 1–40 (cit. on p. 12).
- [ComCom, 2012] OFCOM ComCom. *Results of the auction of all mobile radio frequencies in Switzerland*. Tech. rep. <http://www.news.admin.ch/NSBSubscriber/message/attachments/26004.pdf>. ComCom, OFCOM, 2012 (cit. on p. 29).
- [ComReg, 2021] ComReg. *Multi Band Spectrum Award–Information Memorandum and Draft Regulations*. Tech. rep. <https://www.comreg.ie/media/2021/04/ComReg-2140.pdf>. ComReg, 2021 (cit. on p. 30).

- [Coulom, 2006] Rémi Coulom. “Efficient selectivity and backup operators in Monte-Carlo tree search”. In: *International conference on computers and games*. Springer. 2006, pp. 72–83 (cit. on pp. 79, 81, 83).
- [Coulom, 2008] Rémi Coulom. “Whole-history rating: A Bayesian rating system for players of time-varying strength”. In: *International conference on computers and games*. Springer. 2008, pp. 113–124 (cit. on p. 44).
- [Cowling, 2012a] Peter I Cowling, Edward J Powley, and Daniel Whitehouse. “Information set monte carlo tree search”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 4.2 (2012), pp. 120–143 (cit. on pp. 39, 85, 87, 158, 159, 163, 167).
- [Cowling, 2012b] Peter I Cowling, Colin D Ward, and Edward J Powley. “Ensemble determinization in monte carlo tree search for the imperfect information card game magic: The gathering”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 4.4 (2012), pp. 241–257 (cit. on pp. 86, 88).
- [Cramton, 2001] Peter Cramton and Jesse A Schwartz. “Collusive bidding in the FCC spectrum auctions”. In: *Contributions in Economic Analysis & Policy* 1.1 (2001) (cit. on p. 24).
- [Cramton, 2002] Peter Cramton et al. “Spectrum auctions”. In: *Handbook of telecommunications economics* 1 (2002), pp. 605–639 (cit. on pp. 11, 15, 21, 22, 24, 38, 55).
- [Cramton, 2006] Peter Cramton et al. “Simultaneous ascending auctions”. In: *Combinatorial auctions* (2006), pp. 99–114 (cit. on pp. 21, 23, 32, 55, 94).
- [Cramton, 2007] Peter Cramton, Yoav Shoham, and Richard Steinberg. “An overview of combinatorial auctions”. In: *ACM SIGecom Exchanges* 7.1 (2007), pp. 3–14 (cit. on p. 19).
- [Day, 2012] Robert W Day and Peter Cramton. “Quadratic core-selecting payment rules for combinatorial auctions”. In: *Operations Research* 60.3 (2012), pp. 588–603 (cit. on p. 26).
- [Dr Gerdis Marquardt, 2021] Serafino Abate Dr Gerdis Marquardt Neil Gallagher. *5G spectrum auctions in Europe – policy objectives and auction design*. Tech. rep. Copenhagen Economics, 2021 (cit. on pp. 9, 29).
- [European 5G Observatory, 2018] European 5G Observatory. *Italian 5G spectrum auction*. <https://5gobservatory.eu/italian-5g-spectrum-auction-2/>. European 5G Observatory. 2018 (cit. on pp. 22, 40, 100, 127).

- [European 5G Observatory, 2019] European 5G Observatory. *Japan assigns 5G spectrum to four operators*. <https://5gobservatory.eu/japan-assigns-5g-spectrum-to-four-operators/>. European 5G Observatory. 2019 (cit. on p. 11).
- [FCC, 2020] FCC. *America's 5G Future*. <https://www.fcc.gov/5G>. FCC. 2020 (cit. on p. 9).
- [Fern, 2011] Alan Fern and Paul Lewis. "Ensemble monte-carlo planning: An empirical study". In: *Proceedings of the International Conference on Automated Planning and Scheduling*. Vol. 21. 2011, pp. 58–65 (cit. on p. 86).
- [Foreman, 1996] Peter Foreman and J Keith Murnighan. "Learning to avoid the winner's curse". In: *Organizational Behavior and Human Decision Processes* 67.2 (1996), pp. 170–180 (cit. on p. 14).
- [Frank, 1998] Ian Frank and David Basin. "Search in games with incomplete information: A case study using bridge card play". In: *Artificial Intelligence* 100.1-2 (1998), pp. 87–123 (cit. on p. 86).
- [Gelly, 2007] Sylvain Gelly and David Silver. "Combining online and offline knowledge in UCT". In: *Proceedings of the 24th international conference on Machine learning*. 2007, pp. 273–280 (cit. on p. 81).
- [Gelly, 2011] Sylvain Gelly and David Silver. "Monte-Carlo tree search and rapid action value estimation in computer Go". In: *Artificial Intelligence* 175.11 (2011), pp. 1856–1875 (cit. on p. 81).
- [Goeree, 2003] Jacob K Goeree and Theo Offerman. "Competitive bidding in auctions with private and common values". In: *The Economic Journal* 113.489 (2003), pp. 598–613 (cit. on p. 174).
- [Goeree, 2014] Jacob K Goeree and Yuanchuan Lien. "An equilibrium analysis of the simultaneous ascending auction". In: *Journal of Economic Theory* 153 (2014), pp. 506–533 (cit. on pp. 36, 44, 56, 95, 100, 115, 123, 124).
- [Grimm, 2003] Veronika Grimm, Frank Riedel, and Elmar Wolfstetter. "Low price equilibrium in multi-unit auctions: the GSM spectrum auction in Germany". In: *International journal of industrial organization* 21.10 (2003), pp. 1557–1569 (cit. on p. 23).
- [GSMA, 2014] GSMA. *The Cost of Spectrum Auction Distortions*. Tech. rep. Global System for Mobile Communications, Oct. 2014 (cit. on pp. 17, 19).
- [Gul, 1999] Faruk Gul and Ennio Stacchetti. "Walrasian equilibrium with gross substitutes". In: *Journal of Economic theory* 87.1 (1999), pp. 95–124 (cit. on pp. 35, 49, 50).

- [Harsanyi, 1967] John C Harsanyi. “Games with incomplete information played by “Bayesian” players, I–III Part I. The basic model”. In: *Management science* 14.3 (1967), pp. 159–182 (cit. on p. 154).
- [Harsanyi, 1968] John C Harsanyi. “Games with incomplete information played by “Bayesian” players part II. Bayesian equilibrium points”. In: *Management Science* 14.5 (1968), pp. 320–334 (cit. on p. 153).
- [Hazlett, 1993] Thomas W Hazlett and Robert J Michaels. “The cost of rent-seeking: Evidence from cellular telephone license lotteries”. In: *Southern Economic Journal* (1993), pp. 425–435 (cit. on p. 12).
- [Hoeffding, 1994] Wassily Hoeffding. “Probability inequalities for sums of bounded random variables”. In: *The collected works of Wassily Hoeffding*. Springer, 1994, pp. 409–426 (cit. on p. 107).
- [Infocomm, 2020] Infocomm Media Development Authority. *5G Call for Proposal (“5G CFP 2020”) - 3.5 GHz Spectrum Rights, mmWave Spectrum Rights*. <https://www.imda.gov.sg/regulations-and-licensing-listing/spectrum-management-and-coordination/spectrum-rights-auctions-and-assignment/5G-CFP-2020>. Infocomm Media Development Authority. 2020 (cit. on p. 11).
- [Jain, 1984] Rajendra K Jain, Dah-Ming W Chiu, William R Hawe, et al. “A quantitative measure of fairness and discrimination”. In: *Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA 21* (1984) (cit. on p. 14).
- [Kelso Jr, 1982] Alexander S Kelso Jr and Vincent P Crawford. “Job matching, coalition formation, and gross substitutes”. In: *Econometrica: Journal of the Econometric Society* (1982), pp. 1483–1504 (cit. on p. 50).
- [Knuth, 1975] Donald E Knuth and Ronald W Moore. “An analysis of alpha-beta pruning”. In: *Artificial intelligence* 6.4 (1975), pp. 293–326 (cit. on p. 70).
- [Kocsis, 2006] Levente Kocsis and Csaba Szepesvári. “Bandit based monte-carlo planning”. In: *European conference on machine learning*. Springer. 2006, pp. 282–293 (cit. on pp. 80, 83, 107, 134).
- [Kokott, 2017] Gian-Marco Kokott, Martin Bichler, and Per Paulsen. *Combinatorial first-price auctions: Theory and experiments*. Tech. rep. Working Paper. Technical University of Munich, 2017 (cit. on p. 19).



- [Lasaulce, 2011] Samson Lasaulce and Hamidou Tembine. *Game theory and learning for wireless networks: fundamentals and applications*. Academic Press, 2011 (cit. on pp. 64, 68).
- [Lattimore, 2020] Tor Lattimore and Csaba Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020 (cit. on pp. 76–78, 113).
- [Lee, 2009] Chang-Shing Lee, Mei-Hui Wang, Guillaume Chaslot, Jean-Baptiste Hoock, Arpad Rimmel, Olivier Teytaud, Shang-Rong Tsai, Shun-Chin Hsu, and Tzung-Pei Hong. “The computational intelligence of MoGo revealed in Taiwan’s computer Go tournaments”. In: *IEEE Transactions on Computational Intelligence and AI in games* 1.1 (2009), pp. 73–89 (cit. on p. 83).
- [Lee, 2018a] Jongmin Lee, Geon-Hyeong Kim, Pascal Poupart, and Kee-Eung Kim. “Monte-Carlo tree search for constrained POMDPs”. In: *Advances in Neural Information Processing Systems* 31 (2018) (cit. on p. 132).
- [Lee, 2018b] Jongmin Lee, Geon-hyeong Kim, Pascal Poupart, and Kee-Eung Kim. “Monte-Carlo Tree Search for Constrained MDPs”. In: *IJCAI 2018*. 2018 (cit. on p. 132).
- [Lehmann, 2006] Benny Lehmann, Daniel Lehmann, and Noam Nisan. “Combinatorial auctions with decreasing marginal utilities”. In: *Games and Economic Behavior* 55.2 (2006), pp. 270–296 (cit. on pp. 33, 34).
- [Liu, 2020] Xingchi Liu, Mahsa Derakhshani, Sangarapillai Lambotharan, and Mihaela Van der Schaar. “Risk-Aware Multi-Armed Bandits With Refined Upper Confidence Bounds”. In: *IEEE Signal Processing Letters* 28 (2020), pp. 269–273 (cit. on p. 120).
- [Lorentz, 2016] Richard Lorentz. “Using evaluation functions in Monte-Carlo tree search”. In: *Theoretical computer science* 644 (2016), pp. 106–113 (cit. on p. 83).
- [Luckhart, 1986] Carol Luckhart and Keki B Irani. “An Algorithmic Solution of N-Person Games.” In: *AAAI*. Vol. 86. 1986, pp. 158–162 (cit. on p. 72).
- [Lucking-Reiley, 2000] David Lucking-Reiley. “Vickrey auctions in practice: From nineteenth-century philately to twenty-first-century e-commerce”. In: *Journal of economic perspectives* 14.3 (2000), pp. 183–192 (cit. on pp. 17, 19).
- [Maschler, 2020] Michael Maschler, Shmuel Zamir, and Eilon Solan. *Game theory*. Cambridge University Press, 2020 (cit. on p. 38).
- [McMillan, 1994] John McMillan. “Selling spectrum rights”. In: *Journal of Economic Perspectives* 8.3 (1994), pp. 145–162 (cit. on p. 18).

- [Michie, 1966] Donald Michie. “Game-playing and game-learning automata”. In: *Advances in programming and non-numerical computation*. Elsevier, 1966, pp. 183–200 (cit. on p. 72).
- [Milgrom, 2000] Paul Milgrom. “Putting auction theory to work: The simultaneous ascending auction”. In: *Journal of political economy* 108.2 (2000), pp. 245–272 (cit. on pp. 15, 21, 22, 33, 34, 36, 38, 46–48, 55–57, 94, 95, 113, 116, 138, 179).
- [Milgrom, 2004] Paul Robert Milgrom. *Putting auction theory to work*. Cambridge University Press, 2004 (cit. on pp. 18, 36, 123).
- [Ministère, 2023] Ministère de l'économie des finances et de la souveraineté industrielle et numérique. *Numérique : lancement du programme de recherche « Réseaux du futur » et de la plateforme « France 6G »*. <https://www.economie.gouv.fr/numerique-lancement-programme-recherche-reseaux-futur-plateforme-france-6g>. Ministère de l'économie des finances et de la souveraineté industrielle et numérique. 2023 (cit. on p. 9).
- [Myerson, 1978] Roger B Myerson. “Refinements of the Nash equilibrium concept”. In: *International journal of game theory* 7 (1978), pp. 73–80 (cit. on pp. 13, 37).
- [Nedelec, 2022] Thomas Nedelec, Clément Calauzènes, Noureddine El Karoui, Vianney Perchet, et al. “Learning in repeated auctions”. In: *Foundations and Trends® in Machine Learning* 15.3 (2022), pp. 176–334 (cit. on pp. 13, 153, 154, 174).
- [Neumann, 1947] Ludwig Johann Neumann, Oskar Morgenstern, et al. *Theory of games and economic behavior*. Vol. 60. Princeton university press Princeton, 1947 (cit. on pp. 38, 68).
- [Nijssen, 2012a] JAM Nijssen and Mark HM Winands. “An overview of search techniques in multi-player games”. In: *Computer Games Workshop at ECAI*. 2012, pp. 50–61 (cit. on p. 84).
- [Nijssen, 2012b] Pim Nijssen and Mark HM Winands. “Monte carlo tree search for the hide-and-seeK game scotland yard”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 4.4 (2012), pp. 282–294 (cit. on p. 165).
- [Nijssen, 2013] Joseph Antonius Maria Nijssen. “Monte-Carlo tree search for multi-player games”. PhD thesis. Maastricht University, 2013 (cit. on pp. 84, 107).

- [Ofcom, 2020] Ofcom. *Award of the 700 MHz and 3.6-3.8 GHz spectrum bands*. Tech. rep. [https://www.ofcom.org.uk/{}{}\\_data/assets/pdf{}\\_file/0020/192413/statement-award-700mhz-3.6-3.8ghz-spectrum.pdf](https://www.ofcom.org.uk/{}{}_data/assets/pdf{}_file/0020/192413/statement-award-700mhz-3.6-3.8ghz-spectrum.pdf). Ofcom, 2020 (cit. on p. 22).
- [Osborne, 2004] Martin J Osborne et al. *An introduction to game theory*. Vol. 3. Oxford university press New York, 2004 (cit. on p. 44).
- [Osepayshvili, 2012] Anna Osepayshvili, Michael P Wellman, Daniel Reeves, and Jeffrey K MacKie-Mason. “Self-confirming price prediction for bidding in simultaneous ascending auctions”. In: *arXiv preprint arXiv:1207.1400* (2012) (cit. on p. 55).
- [Pacaud, 2022] Alexandre Pacaud, Marceau Coupechoux, and Aurelien Bechler. “Monte Carlo Tree Search Bidding Strategy for Simultaneous Ascending Auctions”. In: *2022 20th International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt)*. IEEE, 2022, pp. 322–329 (cit. on pp. 55, 94, 124).
- [Pacaud, 2023] Alexandre Pacaud, Aurelien Bechler, and Marceau Coupechoux. “Bidding efficiently in Simultaneous Ascending Auctions with budget and eligibility constraints using Simultaneous Move Monte Carlo Tree Search”. In: *arXiv preprint arXiv:2307.11428* (2023) (cit. on p. 122).
- [Pekeč, 2003] Aleksandar Pekeč and Michael H Rothkopf. “Combinatorial auction design”. In: *Management science* 49.11 (2003), pp. 1485–1503 (cit. on p. 15).
- [Pepels, 2016] Tom Pepels, Tristan Cazenave, and Mark HM Winands. “Sequential halving for partially observable games”. In: *Computer Games: Fourth Workshop on Computer Games, CGW 2015, and the Fourth Workshop on General Intelligence in Game-Playing Agents, GIGA 2015, Held in Conjunction with the 24th International Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 26-27, 2015, Revised Selected Papers 4*. Springer, 2016, pp. 16–29 (cit. on p. 167).
- [Perez Liebana, 2015] Diego Perez Liebana, Jens Dieskau, Martin Hunermund, Sanaz Mostaghim, and Simon Lucas. “Open loop search for general video game playing”. In: *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*. 2015, pp. 337–344 (cit. on p. 96).

- [Perez-Liebana, 2015] Diego Perez-Liebana, Spyridon Samothrakis, Julian Togelius, Tom Schaul, Simon M Lucas, Adrien Couëtoux, Jerry Lee, Chong-U Lim, and Tommy Thompson. “The 2014 general video game playing competition”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 8.3 (2015), pp. 229–243 (cit. on p. 84).
- [Peters, 2006] Michael Peters and Sergei Severinov. “Internet auctions with many traders”. In: *Journal of Economic Theory* 130.1 (2006), pp. 220–245 (cit. on p. 48).
- [Ponsen, 2010] Marc JV Ponsen, Geert Gerritsen, and Guillaume Chaslot. “Integrating Opponent Models with Monte-Carlo Tree Search in Poker.” In: *Interactive Decision Theory and Game Theory* 82 (2010) (cit. on p. 88).
- [Porter, 2006] David Porter and Vernon Smith. “FCC license auction design: A 12-year experiment”. In: *JL Econ. & Pol’y* 3 (2006), p. 63 (cit. on pp. 35, 60).
- [Reeves, 2005] Daniel M Reeves, Michael P Wellman, Jeffrey K MacKie-Mason, and Anna Osepayshvili. “Exploring bidding strategies for market-based scheduling”. In: *Decision Support Systems* 39.1 (2005), pp. 67–85 (cit. on pp. 9, 32, 36, 47, 115, 116, 174).
- [Riedel, 2006] Frank Riedel and Elmar Wolfstetter. “Immediate demand reduction in simultaneous ascending-bid auctions: a uniqueness result”. In: *Economic Theory* 29 (2006), pp. 721–726 (cit. on p. 58).
- [Robles, 2011] David Robles, Philipp Rohlfshagen, and Simon M Lucas. “Learning non-random moves for playing Othello: Improving Monte Carlo tree search”. In: *2011 IEEE Conference on Computational Intelligence and Games (CIG’11)*. IEEE. 2011, pp. 305–312 (cit. on p. 83).
- [Rojiers, 2013] Diederik M Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. “A survey of multi-objective sequential decision-making”. In: *Journal of Artificial Intelligence Research* 48 (2013), pp. 67–113 (cit. on p. 45).
- [Rothkopf, 1990] Michael H Rothkopf, Thomas J Teisberg, and Edward P Kahn. “Why are Vickrey auctions rare?” In: *Journal of Political Economy* 98.1 (1990), pp. 94–109 (cit. on p. 17).
- [Rothkopf, 1998] Michael H Rothkopf, Aleksandar Pekeč, and Ronald M Harstad. “Computationally manageable combinational auctions”. In: *Management science* 44.8 (1998), pp. 1131–1147 (cit. on pp. 15, 16, 19).

- [Rothkopf, 2007] Michael H Rothkopf. “Thirteen reasons why the Vickrey-Clarke-Groves process is not practical”. In: *Operations Research* 55.2 (2007), pp. 191–197 (cit. on p. 20).
- [Russell, 2021] Stuart Russell and Peter Norvig. *Artificial intelligence: a modern approach*. Pearson Education, 2021 (cit. on pp. 40, 70).
- [Schadd, 2009] Maarten PD Schadd, Mark HM Winands, and Jos WHM Uiterwijk. “ChanceProbcut: Forward pruning in chance nodes”. In: *2009 IEEE Symposium on Computational Intelligence and Games*. IEEE. 2009, pp. 178–185 (cit. on p. 72).
- [Schadd, 2011] Maarten PD Schadd and Mark HM Winands. “Best reply search for multiplayer games”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 3.1 (2011), pp. 57–66 (cit. on p. 74).
- [Schaeffer, 2009] Mohammad Shafiei Nathan Sturtevant Jonathan Schaeffer, N Shafiei, et al. “Comparing UCT versus CFR in simultaneous games”. In: *IJCAI Workshop on General Game Playing*. Citeseer. 2009 (cit. on p. 85).
- [Silver, 2016] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. “Mastering the game of Go with deep neural networks and tree search”. In: *nature* 529.7587 (2016), pp. 484–489 (cit. on pp. 38, 44, 75).
- [Silver, 2017] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dhharshan Kumaran, Thore Graepel, et al. “Mastering chess and shogi by self-play with a general reinforcement learning algorithm”. In: *arXiv preprint arXiv:1712.01815* (2017) (cit. on pp. 38, 75).
- [Sturtevant, 2000] Nathan R Sturtevant and Richard E Korf. “On pruning techniques for multi-player games”. In: *AAAI/IAAI* 49 (2000), pp. 201–207 (cit. on p. 73).
- [Sturtevant, 2008] Nathan Sturtevant. “An analysis of UCT in multi-player games”. In: *ICGA Journal* 31.4 (2008), pp. 195–208 (cit. on p. 84).
- [Świechowski, 2023] Maciej Świechowski, Konrad Godlewski, Bartosz Sawicki, and Jacek Mańdziuk. “Monte Carlo tree search: A review of recent modifications and applications”. In: *Artificial Intelligence Review* 56.3 (2023), pp. 2497–2562 (cit. on pp. 39, 81, 135).

- [Szentes, 2003a] Balazs Szentes and Robert W Rosenthal. “Beyond chopsticks: Symmetric equilibria in majority auction games”. In: *Games and Economic Behavior* 45.2 (2003), pp. 278–295 (cit. on pp. 95, 124).
- [Szentes, 2003b] Balázs Szentes and Robert W Rosenthal. “Three-object two-bidder simultaneous auctions: chopsticks and tetrahedra”. In: *Games and Economic Behavior* 44.1 (2003), pp. 114–133 (cit. on pp. 95, 124).
- [Szita, 2009] István Szita, Guillaume Chaslot, and Pieter Spronck. “Monte-carlo tree search in settlers of catan”. In: *Advances in Computer Games*. Springer. 2009, pp. 21–32 (cit. on pp. 81, 82).
- [Tak, 2014] Mandy JW Tak, Marc Lanctot, and Mark HM Winands. “Monte Carlo tree search variants for simultaneous move games”. In: *2014 IEEE Conference on Computational Intelligence and Games*. IEEE. 2014, pp. 1–8 (cit. on pp. 85, 132, 202).
- [Tesauro, 1996] Gerald Tesauro and Gregory Galperin. “On-line policy improvement using Monte-Carlo search”. In: *Advances in Neural Information Processing Systems* 9 (1996) (cit. on p. 76).
- [Teytaud, 2009] Fabien Teytaud and Olivier Teytaud. “Creating an upper-confidence-tree program for Havannah”. In: *Advances in Computer Games*. Springer. 2009, pp. 65–74 (cit. on p. 83).
- [Teytaud, 2011] Olivier Teytaud and Sébastien Flory. “Upper confidence trees with short term partial information”. In: *European Conference on the Applications of Evolutionary Computation*. Springer. 2011, pp. 153–162 (cit. on pp. 85, 160).
- [Thaler, 1988] Richard H Thaler. “Anomalies: The winner’s curse”. In: *Journal of economic perspectives* 2.1 (1988), pp. 191–202 (cit. on pp. 14, 19).
- [Theodoridis, 2020] Alexios Theodoridis and Georgios Chalkiadakis. “Monte carlo tree search for the game of diplomacy”. In: *11th Hellenic Conference on Artificial Intelligence*. 2020, pp. 16–25 (cit. on p. 81).
- [Van den Broeck, 2009] Guy Van den Broeck, Kurt Driessens, and Jan Ramon. “Monte-Carlo tree search in poker using expected reward distributions”. In: *Advances in Machine Learning: First Asian Conference on Machine Learning, ACML 2009, Nanjing, China, November 2-4, 2009. Proceedings 1*. Springer. 2009, pp. 367–381 (cit. on p. 88).

- [Vangerven, 2021] Bart Vangerven, Dries R Goossens, and Frits CR Spieksma. “Using feedback to mitigate coordination and threshold problems in iterative combinatorial auctions”. In: *Business & Information Systems Engineering* 63 (2021), pp. 113–127 (cit. on p. 15).
- [Veness, 2011] Joel Veness, Kee Siong Ng, Marcus Hutter, William Uther, and David Silver. “A monte-carlo aixi approximation”. In: *Journal of Artificial Intelligence Research* 40 (2011), pp. 95–142 (cit. on pp. 84, 96).
- [Vickrey, 1961] William Vickrey. “Counterspeculation, auctions, and competitive sealed tenders”. In: *The Journal of finance* 16.1 (1961), pp. 8–37 (cit. on pp. 17, 18).
- [Waehrer, 1998] Keith Waehrer, Ronald M Harstad, and Michael H Rothkopf. “Auction form preferences of risk-averse bid takers”. In: *The RAND Journal of Economics* (1998), pp. 179–192 (cit. on p. 18).
- [Wang, 2015] Jiao Wang, Tan Zhu, Hongye Li, Chu-Hsuan Hsueh, and I-Chen Wu. “Belief-state Monte-Carlo tree search for Phantom games”. In: *2015 IEEE conference on computational intelligence and games (CIG)*. IEEE. 2015, pp. 267–274 (cit. on p. 87).
- [Weber, 1997] Robert J Weber. “Making more from less: Strategic demand reduction in the FCC spectrum auctions”. In: *Journal of Economics & Management Strategy* 6.3 (1997), pp. 529–548 (cit. on pp. 42, 58, 60).
- [Wellman, 2008] Michael P Wellman, Anna Osepayshvilli, Jeffrey K MacKie-Mason, and Daniel Reeves. *Bidding strategies for simultaneous ascending auctions*. B.E. J. Theoret. Econom, 2008 (cit. on pp. 41, 44–49, 51, 53, 54, 56, 58, 59, 94, 95, 100, 101, 105, 113, 115, 116, 120, 124, 138, 143, 174, 179).
- [Wellman, 2012] Michael P Wellman, Eric Sodomka, and Amy Greenwald. “Self-confirming price prediction strategies for simultaneous one-shot auctions”. In: *arXiv preprint arXiv:1210.4915* (2012) (cit. on p. 54).
- [Whitehouse, 2011] Daniel Whitehouse, Edward J Powley, and Peter I Cowling. “Determinization and information set Monte Carlo tree search for the card game Dou Di Zhu”. In: *2011 IEEE Conference on Computational Intelligence and Games (CIG’11)*. IEEE. 2011, pp. 87–94 (cit. on p. 86).

- [Yuan, 2014] Shuai Yuan, Jun Wang, Bowei Chen, Peter Mason, and Sam Seljan. “An empirical study of reserve price optimisation in real-time bidding”. In: *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2014, pp. 1897–1906 (cit. on p. [174](#)).
- [Zermelo, 1913] Ernst Zermelo. “Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels”. In: *Proceedings of the fifth international congress of mathematicians*. Vol. 2. II, Cambridge UP, Cambridge. 1913, pp. 501–504 (cit. on p. [68](#)).
- [Zheng, 2012] Charles Z Zheng. “Jump bidding and overconcentration in decentralized simultaneous ascending auctions”. In: *Games and Economic Behavior* 76.2 (2012), pp. 648–664 (cit. on pp. [57](#), [59](#), [115](#)).
- [Zobrist, 1990] Albert L Zobrist. “A new hashing method with application for game playing”. In: *ICGA Journal* 13.2 (1990), pp. 69–73 (cit. on pp. [89](#), [137](#)).





## Appendix: Example 1 in d-SAA game

Let  $\Gamma$  be the instance of the d-SAA game in *Example 1* (see Section 2.2.3) which is recalled below.

	$v(\{1\})$	$v(\{2\})$	$v(\{1, 2\})$
Player 1	12	12	12
Player 2	0	0	20

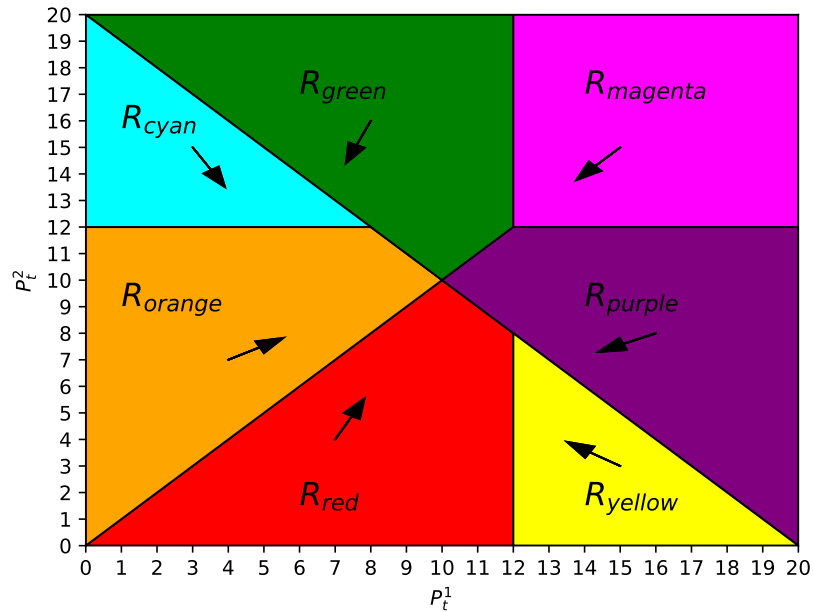
Tab. A.1.: *Example 1*

### A.1 Function $f_\Gamma$

Let  $p = (p_1, p_2)$  be the initial prediction. Let  $f_\Gamma(p) = (f_\Gamma^1(p), f_\Gamma^2(p))$ . We have:

$$f_\Gamma(p) = \begin{cases} (1, 0) & \text{if } p_1 + p_2 \geq 20, p_1 \leq p_2 \text{ and } p_1 < 12 \quad (R_{green}) \\ (0, 1) & \text{if } p_1 + p_2 \geq 20, p_1 > p_2 \text{ and } p_2 < 12 \quad (R_{purple}) \\ (11, 12) & \text{if } p_1 + p_2 < 20 \text{ and } 12 > p_1 > p_2 \quad (R_{red}) \\ (12, 11) & \text{if } p_1 + p_2 < 20 \text{ and } p_1 \leq p_2 < 12 \quad (R_{orange}) \\ (12, 1) & \text{if } p_1 + p_2 < 20 \text{ and } p_2 \geq 12 \quad (R_{cyan}) \\ (1, 12) & \text{if } p_1 + p_2 < 20 \text{ and } p_1 \geq 12 \quad (R_{yellow}) \\ (0, 0) & \text{if } p_1 + p_2 \geq 20 \text{ and } p_1 \geq 12 \text{ and } p_2 \geq 12 \quad (R_{magenta}) \end{cases}$$

$f_\Gamma$  is represented in *Figure A.1*. The arrows correspond to the direction that the sequence  $P_{t+1} = \frac{1}{t+1}f_\Gamma(P_t) + (1 - \frac{1}{t+1})P_t$  will take in each coloured region. Each coloured region in *Figure A.1* will be referred to as  $R_{colour}$ . For example,  $R_{red} = \{(p_1, p_2) | p_1 + p_2 < 20, p_2 < p_1 < 12\}$ .



**Fig. A.1.:** Spatial representation of  $f_\Gamma$  and sequence dynamic in the d-SAA game of *Example 1*

## A.2 Convergence of sequence $P_t$ in the d-SAA game of Example 1

Let's show that the sequence  $P_{t+1} = \frac{1}{t+1}f_\Gamma(P_t) + (1 - \frac{1}{t+1})P_t$  with  $P_0 = (0, 0)$  converges to the frontier prediction of final prices  $O_\Gamma^{min*} = \{(10, 10)\}$ .

The proof is divided into three parts with each a geometrical interpretation:

1. The sequence  $P_t$  is never in  $R_{magenta}$ .
2. The sequence  $P_t^1 - P_t^2$  converges to zero and, thus from a certain iteration onwards,  $P_t$  is never in  $R_{cyan}$  and  $R_{yellow}$ .
3. The sequence  $P_t^1 + P_t^2$  converges to 20 and conclude that  $P_t \xrightarrow[t \rightarrow +\infty]{} (10, 10)$ .

### A.2.1 $P_t$ is never in $R_{magenta}$

*Proof.* Notice that  $\max_{p \in \mathbb{R}_+^2} f_\Gamma^1(p) = \max_{p \in \mathbb{R}_+^2} f_\Gamma^2(p) = 12$ . Then,  $\forall p \in \mathbb{R}_+^2, f_\Gamma^1(p) = 12 \Rightarrow f_\Gamma^2(p) \leq 11$  and  $f_\Gamma^2(p) = 12 \Rightarrow f_\Gamma^1(p) \leq 11$ .

Suppose  $P_t^1 > 11.5$ .

As  $P_t^1 = \frac{1}{t} \sum_{\tau=0}^{t-1} f_\Gamma^1(P_\tau) + \frac{1}{t} P_0^1$  and  $P_0^1 = 0$ , this implies that there is at least  $\lceil \frac{t}{2} \rceil$  values in  $\{0, \dots, t-1\}$  such that  $f_\Gamma^1(P_\tau) = 12$ . Therefore, there is at least  $\lceil \frac{t}{2} \rceil$  values in  $\{0, \dots, t-1\}$ , such that  $f_\Gamma^2(P_\tau) \leq 11$ . Thus, we have:

$$P_t^2 = \frac{1}{t} \sum_{\tau=0}^{t-1} f_\Gamma^2(P_\tau) \leq \frac{1}{t} \left( \sum_{\tau=0}^{\lceil \frac{t}{2} \rceil - 1} 11 + \sum_{\tau=\lceil \frac{t}{2} \rceil}^{t-1} 12 \right) \leq 11.5 \quad (\text{A.1})$$

Symmetrically, if  $P_t^2 > 11.5$  then  $P_t^1 \leq 11.5$ .

This proves that  $\forall t \in \mathbb{N}$ , either  $P_t^1 \leq 11.5$  or  $P_t^2 \leq 11.5$ . Thus, the sequence  $P_t$  is never in  $R_{magenta}$ .

□

### A.2.2 Convergence of $P_t^1 - P_t^2$ to zero

*Proof.* Let's prove that  $U_t = P_t^1 - P_t^2$  converges to zero.

According to the previous result,  $\forall t \in \mathbb{N}$  either  $P_t^1 < 12$  or  $P_t^2 < 12$  which means that  $P_t$  is never in  $R_{magenta}$ . Therefore, if  $P_t^1 > P_t^2$  then  $f_\Gamma^1(P_t) + 12 \geq f_\Gamma^2(P_t) \geq f_\Gamma^1(P_t) + 1$  and if  $P_t^1 \leq P_t^2$  then  $f_\Gamma^2(P_t) + 12 \geq f_\Gamma^1(P_t) \geq f_\Gamma^2(P_t) + 1$ .

We define the sequence  $U_t$  as  $U_t = P_t^1 - P_t^2$  with  $U_0 = 0$ . Thus,  $U_{t+1} = \frac{t}{t+1} U_t + \frac{1}{t+1} (f_\Gamma^1(P_t) - f_\Gamma^2(P_t))$ . Therefore, we have:

- if  $U_t > 0$ :  $\frac{t}{t+1} U_t - \frac{12}{t+1} \leq U_{t+1} \leq \frac{t}{t+1} U_t - \frac{1}{t+1}$
- if  $U_t \leq 0$ :  $\frac{t}{t+1} U_t + \frac{1}{t+1} \leq U_{t+1} \leq \frac{t}{t+1} U_t + \frac{12}{t+1}$

Let's show by induction that  $\forall t \in \mathbb{N}, U_t \in [-\frac{12}{t}, \frac{12}{t}]$ .

1.  $U_1 = -1$  so  $U_1 \in [-12, 12]$

2. Suppose  $U_t \in [-\frac{12}{t}, \frac{12}{t}]$ .

If  $U_t > 0$ , then  $U_t \in ]0, \frac{12}{t}]$  and so, following the inequalities on  $U_{t+1}$ , we have:

$$\frac{-12}{t+1} \leq U_{t+1} \leq \frac{12}{t+1} - \frac{1}{t+1} \leq \frac{12}{t+1} \quad (\text{A.2})$$

Thus, if  $U_t > 0$ , then  $U_{t+1} \in [-\frac{12}{t+1}, \frac{12}{t+1}]$

If  $U_t \leq 0$ , then  $U_t \in [-\frac{12}{t}, 0]$  and so, following the inequalities on  $U_{t+1}$ , we have:

$$\frac{-12}{t+1} \leq \frac{-12}{t+1} + \frac{1}{t+1} \leq U_{t+1} \leq \frac{12}{t+1} \quad (\text{A.3})$$

Thus, if  $U_t \leq 0$ , then  $U_{t+1} \in [-\frac{12}{t+1}, \frac{12}{t+1}]$

Therefore, this proves by induction that  $\forall t \in \mathbb{N}, U_t \in [-\frac{12}{t}, \frac{12}{t}]$ .

From this result, it trivially follows that  $U_t \xrightarrow[t \rightarrow +\infty]{} 0$ .

□

From now on, we will write  $P_t^2 = P_t^1 + \eta_t$  with  $\eta_t \xrightarrow[t \rightarrow +\infty]{} 0$ . In this part, we have proved that the sequence converges to the diagonal  $p_1 = p_2$ . Moreover, combining this result with the result that  $\forall t \in \mathbb{N}$  either  $P_t^1 \leq 11.5$  or  $P_t^2 \leq 11.5$ , we have that  $\forall \epsilon > 0, \exists T \in \mathbb{N}, \forall t \geq T, P_t^1 \leq 11.5 + \epsilon$  and  $P_t^2 \leq 11.5 + \epsilon$ . Therefore, from a certain iteration  $T$  onwards,  $P_t$  can only be in the regions  $R_{green}, R_{purple}, R_{red}$  and  $R_{orange}$ . Thus, we have  $\forall t \geq T$ :

$$U_{t+1} = \begin{cases} \frac{tU_t}{t+1} - \frac{1}{t+1} & \text{if } U_t > 0 \quad (P_t \in \{R_{red}, R_{purple}\}) \\ \frac{tU_t}{t+1} + \frac{1}{t+1} & \text{if } U_t \leq 0 \quad (P_t \in \{R_{orange}, R_{green}\}) \end{cases} \quad (\text{A.4})$$

We can then show that  $\exists t_0 \geq T, \forall t \geq t_0, U_t \in [-\frac{1}{t}, \frac{1}{t}]$  and, hence,  $\eta_t \in [-\frac{1}{t}, \frac{1}{t}]$ . We omit for simplicity the proof of existence of such a  $t_0$  as it is very similar to the following proof.

### A.2.3 Convergence of $P_t$ to (10,10) and geometrical interpretation

*Proof.* As shown previously,  $\exists T \in \mathbb{N}, \forall t \geq T, P_t \in \{R_{green}, R_{purple}, R_{red}, R_{orange}\}$ .

Let fix such a  $T$ . For  $t \geq T$ , we define  $W_t = P_t^1 + P_t^2 - 20$ . We have:

$$\begin{aligned}
 W_{t+1} &= P_{t+1}^1 + P_{t+1}^2 - 20 \\
 &= \frac{f_{\Gamma}^1(P_t) + f_{\Gamma}^2(P_t)}{t+1} + \frac{t(P_t^1 + P_t^2)}{t+1} - 20 \\
 &= \frac{f_{\Gamma}^1(P_t) + f_{\Gamma}^2(P_t)}{t+1} + \frac{tW_t}{t+1} + \frac{20t}{t+1} - 20 \\
 &= \frac{f_{\Gamma}^1(P_t) + f_{\Gamma}^2(P_t)}{t+1} + \frac{tW_t}{t+1} - \frac{20}{t+1}
 \end{aligned} \tag{A.5}$$

Thus,

$$W_{t+1} = \begin{cases} \frac{tW_t}{t+1} + \frac{3}{t+1} & \text{if } W_t < 0 \quad (P_t \in \{R_{red}, R_{orange}\}) \\ \frac{tW_t}{t+1} - \frac{19}{t+1} & \text{if } W_t \geq 0 \quad (P_t \in \{R_{purple}, R_{green}\}) \end{cases} \tag{A.6}$$

**It is now clear that  $W_t \xrightarrow[t \rightarrow +\infty]{} 0$ . Thus,  $P_t^1 \xrightarrow[t \rightarrow +\infty]{} 10$  and  $P_t^2 \xrightarrow[t \rightarrow +\infty]{} 10$ .**

Now, let's show that, from a certain iteration  $t_2 \geq T$ ,  $P_t$  belongs to the diamond defined by the points  $(10 - \frac{10}{t}, 10 - \frac{9}{t})$ ,  $(10 - \frac{9}{t}, 10 - \frac{10}{t})$ ,  $(10 + \frac{2}{t}, 10 + \frac{1}{t})$  and  $(10 + \frac{1}{t}, 10 + \frac{2}{t})$  which progressively converges to the point (10, 10).

**Let's first show that  $\exists t_1 \geq T$  such that  $\forall t \geq t_1, W_t \in [-\frac{19}{t}, \frac{3}{t}]$**

1. First let's show that  $\exists t_1 \geq T, W_{t_1} \in [-\frac{19}{t_1}, \frac{4}{t_1}]$ . We need to look at three different cases. Either  $W_T \in [-\frac{19}{T}, \frac{3}{T}]$ ,  $W_T < -\frac{19}{T}$  or  $W_T > \frac{3}{T}$ .

- a) If  $W_T \in [-\frac{19}{T}, \frac{3}{T}]$ , then the statement is trivially verified.

b) If  $W_T < -\frac{19}{T} < 0$ , let's show that  $\exists t_1 \geq T+1$ ,  $W_{t_1} \in [-\frac{19}{t_1}, \frac{3}{t_1}]$  and  $W_{t_1-1} < -\frac{19}{t_1-1}$ .

First, let's show by contradiction that  $\exists t_1 \geq T+1$ ,  $W_{t_1} \geq -\frac{19}{t_1}$ .

Suppose that  $\forall t \geq T$ ,  $W_t < -\frac{19}{t} < 0$ .

Then, we have have that:

$$\forall t \geq T, W_{t+1} - W_t = \frac{3}{t+1} - \frac{W_t}{t+1} \geq \frac{3}{t+1} \quad (\text{A.7})$$

Therefore,  $\forall t \geq T$ ,  $W_{t+1} - W_T \geq \sum_{i=T}^t \frac{3}{i+1}$ .

As  $\sum_{i=T}^t \frac{3}{i+1} \xrightarrow[t \rightarrow +\infty]{} +\infty$ , this contradicts the fact that  $\forall t \geq T$ ,  $W_t < -\frac{19}{t}$ .

Therefore,  $\exists t_1 \geq T+1$ ,  $W_{t_1} \geq -\frac{19}{t_1}$  and  $W_{t_1-1} < -\frac{19}{t_1-1}$ .

Let's fix such a  $t_1$ .

Let's show by contradiction that  $W_{t_1} \in [-\frac{19}{t_1}, \frac{3}{t_1}]$ .

Suppose  $W_{t_1} \notin [-\frac{19}{t_1}, \frac{3}{t_1}]$ . Then,  $W_{t_1} > \frac{3}{t_1}$ . As  $W_{t_1-1} < -\frac{19}{t_1-1}$ , this implies that  $W_{t_1} - W_{t_1-1} > \frac{3}{t_1} + \frac{19}{t_1-1} > \frac{22}{t_1}$ .

As  $W_{t_1-1} < 0$ ,  $W_{t_1} - W_{t_1-1} = \frac{3}{t_1} - \frac{W_{t_1-1}}{t_1}$ .

Hence,  $W_{t_1-1} \leq -19$ . However,

$$\begin{aligned} \forall W \leq -19, \frac{W(t_1-1)}{t_1} + \frac{3}{t_1} &\leq \frac{-19(t_1-1)}{t_1} + \frac{3}{t_1} \\ &\leq -19 + \frac{22}{t_1} \\ &\leq \frac{3}{t_1} \quad (\text{as } t_1 \geq 1) \end{aligned} \quad (\text{A.8})$$

Therefore, if  $W_{t_1-1} \leq -19$ , then  $W_{t_1} \leq \frac{3}{t_1}$ . This contradicts the fact that  $W_{t_1} > \frac{3}{t_1}$  so  $\exists t_1 \geq T+1$ ,  $W_{t_1} \in [-\frac{19}{t_1}, \frac{3}{t_1}]$  and  $W_{t_1-1} < -\frac{19}{t_1-1}$ . The statement is verified.

c) Similarly to the upper proof, if  $W_T > \frac{4}{T} \geq 0$ ,  $\exists t_1 \geq T + 1$ ,  $W_{t_1} \in [-\frac{19}{t_1}, \frac{3}{t_1}]$  and  $W_{t_1-1} > \frac{3}{t_1-1}$ .

We have shown that  $\exists t_1 \geq T$ ,  $W_{t_1} \in [-\frac{19}{t_1}, \frac{3}{t_1}]$ .

2. Let's show that if  $W_t \in [-\frac{19}{t}, \frac{3}{t}]$  then  $W_{t+1} \in [-\frac{19}{t+1}, \frac{3}{t+1}]$ .

Suppose  $W_t \in [-\frac{19}{t}, \frac{3}{t}]$ .

If  $-\frac{19}{t} \leq W_t < 0$ , we have:

$$-\frac{19}{t+1} \leq \frac{3}{t+1} - \frac{19}{t+1} \leq W_{t+1} \leq \frac{3}{t+1} \quad (\text{A.9})$$

So  $W_{t+1} \in [-\frac{19}{t+1}, \frac{3}{t+1}]$ .

If  $0 \leq W_t \leq \frac{3}{t}$ , we have:

$$-\frac{19}{t+1} \leq W_{t+1} \leq -\frac{19}{t+1} + \frac{3}{t+1} \leq \frac{3}{t+1} \quad (\text{A.10})$$

So  $W_{t+1} \in [-\frac{19}{t+1}, \frac{3}{t+1}]$ .

We have shown by induction that  $\exists t_1 \geq T$ ,  $\forall t \geq t_1$ ,  $W_t \in [-\frac{19}{t}, \frac{3}{t}]$ . Moreover, we have shown that there exist  $\exists t_0 \geq T$ ,  $\forall t \geq t_0$ ,  $U_t \in [-\frac{1}{t}, \frac{1}{t}]$ .

Let fix such a  $t_0$  and  $t_1$ . Let  $t_2 = \max(t_0, t_1)$ .  $\forall t \geq t_2$ , we have:

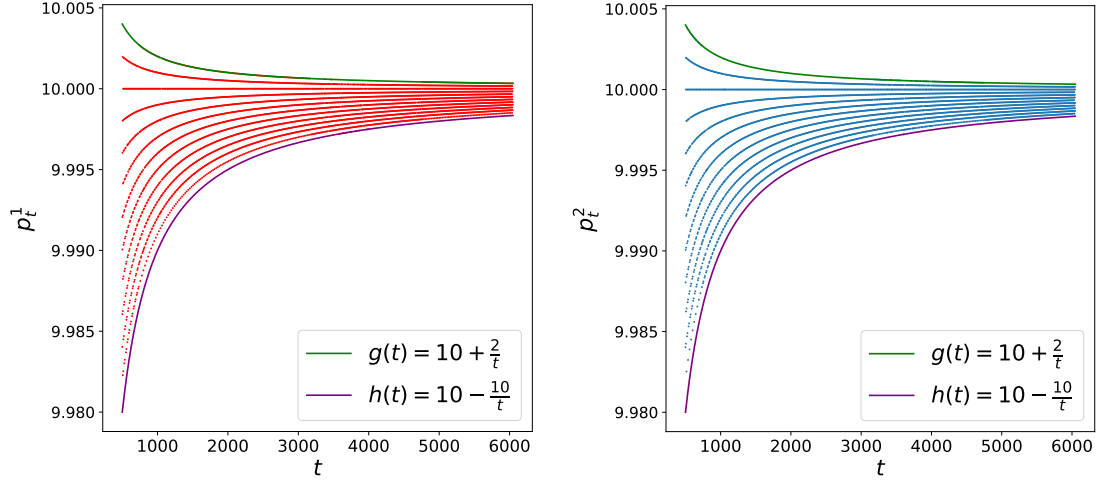
$$\begin{aligned} & -\frac{19}{t} \leq W_t \leq \frac{3}{t} \\ \Leftrightarrow & -\frac{19}{t} \leq 2P_t^1 + \eta_t - 20 \leq \frac{3}{t} \\ \Leftrightarrow & -\frac{19}{2t} + 10 - \frac{\eta_t}{2} \leq P_t^1 \leq \frac{3}{2t} + 10 - \frac{\eta_t}{2} \end{aligned} \quad (\text{A.11})$$



As  $\eta_t \in [-\frac{1}{t}, \frac{1}{t}]$ , we have that:

$$\forall t \geq t_2, 10 - \frac{10}{t} \leq P_t^1 \leq 10 + \frac{2}{t} \quad \text{and} \quad 10 - \frac{10}{t} \leq P_t^2 \leq 10 + \frac{2}{t} \quad (\text{A.12})$$

The sequence  $P_t^1$  and  $P_t^2$  along with their upper-bound  $g(t) = 10 + \frac{2}{t}$  and lower-bound  $h(t) = 10 - \frac{10}{t}$  from a certain iteration  $i_0$  are illustrated in Figure A.2.



**Fig. A.2.:** Convergence rate of  $P_t^1$  and  $P_t^2$  in the d-SAA game of *Example 1*.

Geometrically, we have first shown, through  $U_t \in [-\frac{1}{t}, \frac{1}{t}]$ , that  $P_t$  was between the line  $p_1 - p_2 = -\frac{1}{t}$  and  $p_1 - p_2 = \frac{1}{t}$  which are parallel to the line  $p_1 = p_2$ . Then, in this last part, through  $W_t \in [-\frac{19}{t}, \frac{3}{t}]$ , we showed that  $P_t$  from a certain iteration  $i_0$  is also between the lines  $p_1 + p_2 = 20 - \frac{19}{t}$  and  $p_1 + p_2 = 20 + \frac{3}{t}$  which are parallel to the line  $p_1 + p_2 = 20$ . Using both of these information, we know that  $P_t$  belongs to the diamond defined by the points  $(10 - \frac{10}{t}, 10 - \frac{9}{t})$ ,  $(10 - \frac{9}{t}, 10 - \frac{10}{t})$ ,  $(10 + \frac{2}{t}, 10 + \frac{1}{t})$  and  $(10 + \frac{1}{t}, 10 + \frac{2}{t})$  which progressively converges to the point  $(10, 10)$ .

□

## Appendix: Proof of properties 4.2 and 4.3

In this part, we are going to prove *Property 4.2* and *4.3*. More precisely, to prove *Property 4.2*, we will show that, for any scenario which might occur in a d-SAA game with two goods between 2 players with super-additive value functions, the frontier prediction of final prices is reduced to a singleton. To prove *Property 4.3*, we will show in parallel that sequence  $P_t$  converges for each scenario to the only element of the frontier prediction of final prices. All scenarios are covered below.

We will omit the cases where  $\exists i \in \{1, 2\}, \exists j \in \{1, 2\}, v_i(x_j) = 0$ .

### 1. First disposition:

$$\begin{cases} v_1(x_1) \leq v_2(x_1) \\ v_1(x_2) \leq v_2(x_2) \\ v_1(x_1 + x_2) \leq v_2(x_1 + x_2) \end{cases} \quad (\text{B.1})$$

*Proof.* **Set of smallest overestimated final prices:**

The following result is shared amongst all cases of the first disposition. In this proof, we are going to calculate the set of smallest overestimated final prices  $O_\Gamma^{min}$ . To do so, we will first start by calculating the set of overestimated prices  $O_\Gamma$ .

We will show that  $O_\Gamma = \{p, v_1(x_1 + x_2) \leq p_1 + p_2, v_1(x_1) \leq p_1, v_1(x_2) \leq p_2\}$  in two steps:

- If  $p \in O_\Gamma \implies p_1 \geq v_1(x_1), p_2 \geq v_1(x_2), p_1 + p_2 \geq v_1(x_1 + x_2)$
- If  $p_1 \geq v_1(x_1), p_2 \geq v_1(x_2), p_1 + p_2 \geq v_1(x_1 + x_2) \implies p \in O_\Gamma$

As  $v_1(x_1) \leq v_2(x_1)$ ,  $\forall p \in \mathbb{R}_+^2$  such that  $p_1 < v_1(x_1)$ , then  $f_\Gamma^1(p) \geq v_1(x_1)$  as the acquisition of item 1 for a price below  $v_1(x_1)$  will always benefit both players regardless of the price of item 2. Symmetrically, as  $v_1(x_2) \leq v_2(x_2)$ ,  $\forall p \in \mathbb{R}_+^2$  such that  $p_2 < v_1(x_2)$ , then  $f_\Gamma^2(p) \geq v_1(x_2)$ . Thus,  $p \in O_\Gamma \implies p_1 \geq v_1(x_1)$  and  $p_2 \geq v_1(x_2)$ .

Moreover,  $\forall p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 < v_1(x_1 + x_2)$ ,  $p_1 \geq v_1(x_1)$  and  $p_2 \geq v_1(x_2)$ , we have  $f_\Gamma^1(p) > p_1$  or  $f_\Gamma^2(p) > p_2$ . Player 1 will necessarily bid on both items until either player 2 stops bidding or the entirety of her complementarity surplus has been consumed. As  $v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , player 2 will either bid on both items until the entirety of her complementarity surplus has been consumed and, thus, the final price of item 1 (respectively item 2) will be greater than  $p_1$  (respectively  $p_2$ ) or will decide to only bid on one item and the final price of this item will be greater than its initial prediction. Thus,  $p \in O_\Gamma \implies p_1 \geq v_1(x_1), p_2 \geq v_1(x_2), p_1 + p_2 \geq v_1(x_1 + x_2)$ .

The first step is verified

Conversely,  $\forall p \in \mathbb{R}_+^2$  such that  $p_1 \geq v_1(x_1), p_2 \geq v_1(x_2), p_1 + p_2 \geq v_1(x_1 + x_2)$ , we have  $f_\Gamma^1(p) = f_\Gamma^2(p) = 0$  as player 1 predicts getting a negative utility if he eventually acquires any object and, hence, drops out of the auction immediately. Thus,  $p \in O_\Gamma$ .

The second step is verified.

Therefore,

$$O_\Gamma = \{p, v_1(x_1 + x_2) \leq p_1 + p_2, v_1(x_1) \leq p_1, v_1(x_2) \leq p_2\} \quad (\text{B.2})$$

As  $O_\Gamma^{\min} = \operatorname{argmin}_{p \in \bar{O}_\Gamma} \|p\|_1$ , we have:

$$O_\Gamma^{\min} = \{p, v_1(x_1 + x_2) = p_1 + p_2, v_1(x_1) \leq p_1, v_1(x_2) \leq p_2\} \quad (\text{B.3})$$

□

a) **First case:**

$$* \frac{v_1(x_1 + x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_1)$$

$$* \frac{v_1(x_1 + x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_2)$$

i. **First sub-case:**

$$* v_1(x_1) > \frac{v_1(x_1+x_2)}{2} \text{ (or respectively } v_1(x_2) > \frac{v_1(x_1+x_2)}{2}\text{)}$$

**Proof. Frontier prediction of final prices:**

In this proof, we are going to calculate the frontier prediction of final prices and, thus, calculate the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_\Gamma|_{O_\Gamma^{min}}$ , the restriction of  $g_\Gamma$  to  $O_\Gamma^{min}$ , for any  $\gamma > 0$ .

We can rewrite  $O_\Gamma^{min}$  as  $I_\alpha = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2); v_1(x_1 + x_2) - v_1(x_1)]\}$ . From there, we are going to calculate the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_\Gamma|_{O_\Gamma^{min}}$  for any  $\gamma > 0$ . To do so, we need to obtain the closed-form of  $f_\Gamma$  for every slight underestimation of elements in  $I_\alpha$ . We decompose  $I_\alpha$  in 3 disjoint sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2); A]\}$
- $I_\alpha^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [A, v_1(x_1 + x_2) - v_1(x_1)]\}$
- $I_\alpha^3 = \{(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$

with  $A = \max(v_1(x_2), v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)))$  such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3$ . We will calculate the infimum of  $g_\Gamma$  on each of these three sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- If  $v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)) \leq v_1(x_2)$ , i.e.,  $A = v_1(x_2)$ , then  $I_\alpha^1 = \emptyset$ . Otherwise, we have  $v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)) > v_1(x_2)$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p \in I_\alpha^1$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_2)$  and, thus, player 2 won't bid on item 1. The final price of item 1 is then 0. However, player 1 predicts that item 1 will be sold for  $p_1 - \eta_1 > v_1(x_1)$  and, hence, she thinks that she is exposed of  $p_1 - \eta_1 - v_1(x_1)$ . It is then beneficial for player 1 to acquire item 2 if its price is below  $(v_1(x_1 + x_2) - (p_1 - \eta_1)) + (p_1 - \eta_1 - v_1(x_1))$ . The first term comes from the surplus of acquiring item 2 given player 1 has acquired item 1 and the second term comes

from its actual predicted exposure. Player 2 will however continue to bid on item 2 until its price reaches  $v_2(x_2)$ . Thus, if player 1 stops bidding on item 2 first, then its final price is  $v_1(x_1 + x_2) - v_1(x_1)$ . If player 2 stops bidding on item 2 first, then its final price is  $v_2(x_2)$ . We set  $R_2 = \min(v_2(x_2), v_1(x_1 + x_2) - v_1(x_1))$  and, thus, the final price of item 2 is  $R_2$ . More formally, we have  $\forall p \in I_\alpha^1, \exists \eta > 0$  such that  $\forall (\eta_1, \eta_2) \in (0, \eta], f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = 0$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = R_2$  with  $R_2 = \min(v_2(x_2), v_1(x_1 + x_2) - v_1(x_1))$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in [v_1(x_2); v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2))]$ ,

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_1)^2 + (R_2 - p_2)^2 \\ &= (v_1(x_1 + x_2) - \alpha)^2 + (R_2 - \alpha)^2 \end{aligned} \quad (\text{B.4})$$

Thus,  $\forall p \in I_\alpha^1, g_\Gamma(p) \geq (v_2(x_1 + x_2) - v_2(x_2))^2 > 0$ .

- Suppose  $p \in I_\alpha^2$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_1(x_1)$ . As  $\frac{v_1(x_1 + x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $\frac{v_1(x_1 + x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , player 2 will always bid on both items if their price is below  $\frac{v_1(x_1 + x_2)}{2}$ . As  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 also always bids on both items. Thus, the price of each item rises symmetrically. As  $p_2 + \eta_1 < \frac{v_1(x_1 + x_2)}{2} < p_1 - \eta_1$ , the complementarity surplus of player 1 will be entirely consumed when the price of item 2 reaches  $p_2 + \eta_1$ . Player 1 will then drop out of the auction and the final price of each item will be  $p_2 + \eta_1$ . More formally,  $\forall p \in I_\alpha^2, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta], f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in [A; v_1(x_1 + x_2) - v_1(x_1)]$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \quad (\text{B.5})$$

Thus,  $\forall p \in I_\alpha^2, g_\Gamma(p) > 0$ .

- Suppose  $p \in I_\alpha^3$  and, thus,  $p = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction

bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ . As  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_1)$ ,  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_2)$  and  $v_1(x_1+x_2) \leq v_2(x_1+x_2)$ , player 2 will always bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $p_1 - \eta_1 < v_1(x_1)$  and  $p_2 - \eta_2 < v_1(x_1+x_2) - v_1(x_1) < \frac{v_1(x_1+x_2)}{2} < v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus has been entirely consumed by the rise in price of item 2. This happens when the price of item 2 reaches  $p_2$ . Thus, player 2 acquires item 2 at price  $p_2$ . Then, player 1 will only bid on item 1 until its price reaches her stand-alone value  $p_1 = v_1(x_1)$ . As  $v_1(x_1) \leq v_2(x_1)$ , player 2 will also continue to bid on item 1 and, then, acquire it for a final price of  $p_1$ . More formally, for  $p = (v_1(x_1), v_1(x_1+x_2) - v_1(x_1))$ ,  $\exists \eta > 0$  such that  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = p_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2$ . Thus, we have:

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = 0 \quad (\text{B.6})$$

Thus,  $g_\Gamma(v_1(x_1), v_1(x_1+x_2) - v_1(x_1)) = 0$ .

We have shown that:

- If  $I_\alpha^1 \neq \emptyset$ ,  $g_\Gamma|_{O_\Gamma^{min}}$  is lower-bounded by a strictly positive term on  $I_\alpha^1$ .
- $\forall p \in I_\alpha^2 = \{(v_1(x_1+x_2) - \alpha, \alpha), \alpha \in [A, v_1(x_1+x_2) - v_1(x_1)]\}$ ,  $g_\Gamma(p) > 0$ .
- $g_\Gamma(v_1(x_1), v_1(x_1+x_2) - v_1(x_1)) = 0$

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1), v_1(x_1+x_2) - v_1(x_1))\}$ .

□

*Proof. Sequence convergence:*

Let's first show that:

$$\begin{aligned} p &\in [0, v_1(x_1)) \times [0, v_1(x_1 + x_2) - v_1(x_1)) \\ \implies f_\Gamma(p_1, p_2) &= (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) \end{aligned} \quad (\text{B.7})$$

Suppose  $p \in [0, v_1(x_1)) \times [0, v_1(x_1 + x_2) - v_1(x_1))$ . As  $p_1 < v_2(x_1)$ ,  $p_2 < v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_1)$  and  $v_1(x_1 + x_2) - v_1(x_1) < v_2(x_1)$ , player 2 will bid on both items if their price is below  $v_1(x_1 + x_2) - v_1(x_1)$ . As  $p_1 < v_1(x_1)$  and  $p_2 < v_1(x_1 + x_2) - v_1(x_1) < v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed by item 2. Thus, the price of each item rises symmetrically. Player 1 will stop bidding on item 2 when its price reaches  $v_1(x_1 + x_2) - v_1(x_1)$  and will continue to bid on item 1 until its price reaches  $v_1(x_1)$ . Player 2 will then acquire item 2 for  $v_1(x_1 + x_2) - v_1(x_1)$  and will continue bidding on item 1 until its price reaches  $v_2(x_1 + x_2) - v_2(x_2)$ . As  $v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$ , player 2 acquires item 1 for  $v_1(x_1)$ . Thus,  $f_\Gamma(p_1, p_2) = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ .

As  $P_0 = (0, 0) < (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$  component-wise and  $P_{t+1} = \frac{1}{t+1} f_\Gamma(P_t) + (1 - \frac{1}{t+1}) P_t$ , by a simple induction we have that  $\forall t \in \mathbb{R}_+$ ,  $P_t \in [0, v_1(x_1)) \times [0, v_1(x_1 + x_2) - v_1(x_1))$  and, thus,  $\forall t \in \mathbb{R}_+$ ,  $f_\Gamma(P_t) = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ .

The above results directly imply that  $P_t \xrightarrow[t \rightarrow +\infty]{} (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

## ii. Second sub-case:

$$* \quad v_1(x_1) \leq \frac{v_1(x_1 + x_2)}{2} \text{ and } v_1(x_2) \leq \frac{v_1(x_1 + x_2)}{2}$$

*Proof.* **Frontier prediction of final prices:**

We decompose  $I_\alpha$  in 5 disjoint sets

$$\bullet \quad I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2); A)\}$$

- $I_\alpha^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [A, \frac{v_1(x_1+x_2)}{2}]\}$
- $I_\alpha^3 = \{(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})\}$
- $I_\alpha^4 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (\frac{v_1(x_1+x_2)}{2}, B)\}$
- $I_\alpha^5 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (B, v_1(x_1 + x_2) - v_1(x_1))\}$

with  $A = \max(v_1(x_2), v_1(x_1+x_2) - (v_2(x_1+x_2) - v_2(x_2)))$  and  $B = \min(v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1))$  such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3 \cup I_\alpha^4 \cup I_\alpha^5$ . We will calculate the infimum of  $g_\Gamma$  on each of these five sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- If  $v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)) \leq v_1(x_2)$ , i.e.,  $A = v_1(x_2)$ , then  $I_\alpha^1 = \emptyset$ . Otherwise, we have  $v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)) > v_1(x_2)$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p \in I_\alpha^1$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_2)$  and, thus, player 2 won't bid on item 1. The final price of item 1 is then 0. However, player 1 predicts that item 1 will be sold for  $p_1 - \eta_1 > v_1(x_1)$  and, hence, she thinks that she is exposed of  $p_1 - \eta_1 - v_1(x_1)$ . It is then beneficial for player 1 to acquire item 2 if its price is below  $(v_1(x_1 + x_2) - (p_1 - \eta_1)) + (p_1 - \eta_1 - v_1(x_1))$ . The first term comes from the surplus of acquiring item 2 given player 1 has acquired item 1 and the second term comes from its actual predicted exposure. Player 2 will however continue to bid on item 2 until its price reaches  $v_2(x_2)$ . Thus, if player 1 stops bidding on item 2 first, then its final price is  $v_1(x_1 + x_2) - v_1(x_1)$ . If player 2 stops bidding on item 2 first, then its final price is  $v_2(x_2)$ . We set  $R_2 = \min(v_2(x_2), v_1(x_1 + x_2) - v_1(x_1))$  and, thus, the final price of item 2 is  $R_2$ . More formally, we have  $\forall p \in I_\alpha^1, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta], f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = 0$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = R_2$  with  $R_2 = \min(v_2(x_2), v_1(x_1 + x_2) - v_1(x_1))$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in [v_1(x_2); v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2))]$ ,

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_1)^2 + (R_2 - p_2)^2 \\ &= (v_1(x_1 + x_2) - \alpha)^2 + (R_2 - \alpha)^2 \end{aligned} \quad (B.8)$$

Thus,  $\forall p \in I_\alpha^1, g_\Gamma(p) \geq (v_2(x_1 + x_2) - v_2(x_2))^2 > 0$ .



- Suppose  $p \in I_\alpha^2$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > \frac{v_1(x_1+x_2)}{2}$ . As  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_1)$ ,  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_2)$  and  $v_1(x_1+x_2) \leq v_2(x_1+x_2)$ , player 2 will always bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $p_1 - \eta_1 < v_1(x_1+x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1+x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1+x_2)$ , player 1 also always bids on both items. Thus, the price of each item rises symmetrically. As  $p_2 + \eta_1 < \frac{v_1(x_1+x_2)}{2}$  and  $p_1 - \eta_1 > \frac{v_1(x_1+x_2)}{2}$ , the complementarity surplus of player 1 will be entirely consumed when the price of item 2 reaches  $p_2 + \eta_1$ . Player 1 will then drop out of the auction and the final price of each item will be  $p_2 + \eta_1$ . More formally,  $\forall p \in I_\alpha^2, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ . Thus, for  $p = (v_1(x_1+x_2) - \alpha, \alpha)$  with  $\alpha \in [A; \frac{v_1(x_1+x_2)}{2})$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1+x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \quad (\text{B.9})$$

Thus,  $\forall p \in I_\alpha^2, g_\Gamma(p) > 0$ .

- Suppose  $p \in I_\alpha^3$  and, thus,  $p = (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ . As  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_1)$ ,  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_2)$  and  $v_1(x_1+x_2) \leq v_2(x_1+x_2)$ , player 2 will always bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $v_1(x_1+x_2) - v_1(x_1) \geq \frac{v_1(x_1+x_2)}{2}$  and  $v_1(x_1+x_2) - v_1(x_2) \geq \frac{v_1(x_1+x_2)}{2}$ , player 1 will always bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $p_1 - \eta_1 < \frac{v_1(x_1+x_2)}{2}$  and  $p_2 - \eta_2 < \frac{v_1(x_1+x_2)}{2}$ , both players will always bid on both items and, thus, the price of each item will rise symmetrically until it reaches  $\frac{v_1(x_1+x_2)}{2}$ . Then, player 1 will drop out of the auction. Player 2 acquires each item at a price  $p_1 = p_2 = \frac{v_1(x_1+x_2)}{2}$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_1 = p_2$ . Thus, we have

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = 0 \quad (\text{B.10})$$

Thus,  $g_{\Gamma}(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2}) = 0$ .

- Suppose  $p \in I_{\alpha}^4$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > \frac{v_1(x_1+x_2)}{2}$ . As  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_1)$ ,  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_2)$  and  $v_1(x_1+x_2) \leq v_2(x_1+x_2)$ , player 2 will always bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $p_1 - \eta_1 < v_1(x_1+x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1+x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1+x_2)$ , player 1 also always bids on both items. Thus, the price of each item rises symmetrically. As  $p_1 + \eta_2 < \frac{v_1(x_1+x_2)}{2}$  and  $p_2 - \eta_2 > \frac{v_1(x_1+x_2)}{2}$ , the complementarity surplus of player 1 will be entirely consumed when the price of item 1 reaches  $p_1 + \eta_2$ . Player 1 will then drop out of the auction and the final price of each item will be  $p_1 + \eta_2$ . More formally,  $\forall p \in I_{\alpha}^4, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_{\Gamma}^1(p_1 - \eta_1, p_2 - \eta_2) = f_{\Gamma}^2(p_1 - \eta_1, p_2 - \eta_2) = p_1 + \eta_2$ . Thus, for  $p = (v_1(x_1+x_2) - \alpha, \alpha)$  with  $\alpha \in (\frac{v_1(x_1+x_2)}{2}, B]$ , we have:

$$\begin{aligned} \|f_{\Gamma}(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_1 + \eta_2 - p_1)^2 + (p_1 + \eta_2 - p_2)^2 \\ &= (v_1(x_1+x_2) - 2\alpha + \eta_2)^2 + \eta_2^2 \end{aligned} \quad (\text{B.11})$$

Thus,  $\forall p \in I_{\alpha}^4, g_{\Gamma}(p) > 0$ .

- If  $v_2(x_1+x_2) - v_2(x_1) \geq v_1(x_1+x_2) - v_1(x_1)$ , i.e.,  $B = v_1(x_1+x_2) - v_1(x_1)$ , then  $I_{\alpha}^5 = \emptyset$ . Otherwise, we have  $v_2(x_1+x_2) - v_2(x_1) < v_1(x_1+x_2) - v_1(x_1)$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p \in I_{\alpha}^5$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > v_2(x_1+x_2) - v_2(x_1)$  and, thus, player 2 won't bid on item 2. The final price of item 2 is then 0. However, player 1 predicts that item 2 will be sold for  $p_2 - \eta_2 > v_1(x_2)$  and, hence, she thinks that she is exposed of  $p_2 - \eta_2 - v_1(x_2)$ . It is then beneficial for player 1 to acquire item 1 if its price is below  $(v_1(x_1+x_2) - (p_2 - \eta_2)) + (p_2 - \eta_2 - v_1(x_2))$ . The first term comes from the surplus of acquiring item 1 given player 1 has acquired item 2 and the second term comes from its actual predicted exposure. Player 2 will however continue to bid on item 1 until its price reaches  $v_2(x_1)$ . Thus, if player 1 stops bidding on item 1 first, then its final price is  $v_1(x_1+x_2) - v_1(x_2)$ . If player 2 stops bidding on item 1 first,

then its final price is  $v_2(x_1)$ . We set  $R_1 = \min(v_2(x_1), v_1(x_1+x_2) - v_1(x_2))$  and, thus, the final price of item 1 is  $R_1$ . More formally, we have  $\forall p \in I_\alpha^5$ ,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = R_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = 0$  with  $R_1 = \min(v_2(x_1), v_1(x_1+x_2) - v_1(x_2))$ . Thus, for  $p = (v_1(x_1+x_2) - \alpha, \alpha)$  with  $\alpha \in (v_2(x_1+x_2) - v_2(x_1), v_1(x_1+x_2) - v_1(x_1))$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (R_1 - p_1)^2 + (p_2)^2 \\ &= (R_1 + \alpha - v_1(x_1+x_2))^2 + (\alpha)^2 \end{aligned} \quad (\text{B.12})$$

Thus,  $\forall p \in I_\alpha^5$ ,  $g_\Gamma(p) \geq (v_2(x_1+x_2) - v_2(x_1))^2 > 0$ .

We have shown that:

- If  $I_\alpha^1 \neq \emptyset$ ,  $g_\Gamma|_{O_\Gamma^{min}}$  is lower-bounded by a strictly positive term on  $I_\alpha^1$ .
- $\forall p \in I_\alpha^2 = \{(v_1(x_1+x_2) - \alpha, \alpha), \alpha \in [A, \frac{v_1(x_1+x_2)}{2}]\}$ ,  $g_\Gamma(p) > 0$ .
- $g_\Gamma(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2}) = 0$
- $\forall p \in I_\alpha^4 = \{(v_1(x_1+x_2) - \alpha, \alpha), \alpha \in (\frac{v_1(x_1+x_2)}{2}, B]\}$ ,  $g_\Gamma(p) > 0$ .
- If  $I_\alpha^5 \neq \emptyset$ ,  $g_\Gamma|_{O_\Gamma^{min}}$  is lower-bounded by a strictly positive term on  $I_\alpha^5$ .

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})\}$ .  $\square$

*Proof. Sequence convergence:*

Let's first show that:

$$\begin{aligned} p &\in [0, \frac{v_1(x_1+x_2)}{2}) \times [0, \frac{v_1(x_1+x_2)}{2}) \\ \implies f_\Gamma(p_1, p_2) &= (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2}) \end{aligned} \quad (\text{B.13})$$

Suppose  $p \in [0, \frac{v_1(x_1+x_2)}{2}) \times [0, \frac{v_1(x_1+x_2)}{2})$ . As  $p_1 < \frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_2)$ ,  $p_2 < \frac{v_1(x_1+x_2)}{2} \leq v_2(x_1+x_2) - v_2(x_1)$  and  $p_1 + p_2 < v_1(x_1+x_2) \leq$

$v_2(x_1 + x_2)$ , player 2 will bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $p_1 < \frac{v_1(x_1+x_2)}{2} \leq v_1(x_1 + x_2) - v_1(x_2)$  and  $p_2 < \frac{v_1(x_1+x_2)}{2} \leq v_1(x_1 + x_2) - v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed. Thus, the price of each item rises symmetrically. Player 1 will stop bidding on both items when their price reaches  $\frac{v_1(x_1+x_2)}{2}$ . Player 2 will then acquire each item for  $\frac{v_1(x_1+x_2)}{2}$ . Thus,  $f_\Gamma(p_1, p_2) = (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$ .

As  $P_0 = (0, 0) < (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$  component-wise and  $P_{t+1} = \frac{1}{t+1} f_\Gamma(P_t) + (1 - \frac{1}{t+1})P_t$ , by a simple induction we have that  $\forall t \in \mathbb{R}_+, P_t \in [0, \frac{v_1(x_1+x_2)}{2}] \times [0, \frac{v_1(x_1+x_2)}{2}]$  and, thus,  $\forall t \in \mathbb{R}_+, f_\Gamma(P_t) = (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$ .

The above results directly imply that  $P_t \xrightarrow[t \rightarrow +\infty]{} (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

## b) Second case:

$$* \frac{v_1(x_1+x_2)}{2} > v_2(x_1+x_2) - v_2(x_1) \text{ (respectively } \frac{v_1(x_1+x_2)}{2} > v_2(x_1+x_2) - v_2(x_2))$$

### i. First sub-case:

$$* v_2(x_1 + x_2) - v_2(x_1) \geq v_1(x_1 + x_2) - v_1(x_1) \text{ (or respectively } v_2(x_1 + x_2) - v_2(x_2) \geq v_1(x_1 + x_2) - v_1(x_2))$$

*Proof. Frontier prediction of final prices:*

We decompose  $I_\alpha$  in 3 disjoint sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2); A]\}$
- $I_\alpha^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [A, v_1(x_1 + x_2) - v_1(x_1)]\}$
- $I_\alpha^3 = \{(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$

with  $A = \max(v_1(x_2), v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)))$  such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3$ . We will calculate the infimum of  $g_\Gamma$  on each of these three sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- If  $v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)) \leq v_1(x_2)$ , i.e.,  $A = v_1(x_2)$ , then  $I_\alpha^1 = \emptyset$ . Otherwise, we have  $v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)) > v_1(x_2)$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p \in I_\alpha^1$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_2)$  and, thus, player 2 won't bid on item 1. The final price of item 1 is then 0. However, player 1 predicts that item 1 will be sold for  $p_1 - \eta_1 > v_1(x_1)$  and, hence, she thinks that she is exposed of  $p_1 - \eta_1 - v_1(x_1)$ . It is then beneficial for player 1 to acquire item 2 if its price is below  $(v_1(x_1 + x_2) - (p_1 - \eta_1)) + (p_1 - \eta_1 - v_1(x_1))$ . The first term comes from the surplus of acquiring item 2 given player 1 has acquired item 1 and the second term comes from its actual predicted exposure. Player 2 will however continue to bid on item 2 until its price reaches  $v_2(x_2)$ . Thus, if player 1 stops bidding on item 2 first, then its final price is  $v_1(x_1 + x_2) - v_1(x_1)$ . If player 2 stops bidding on item 2 first, then its final price is  $v_2(x_2)$ . We set  $R_2 = \min(v_2(x_2), v_1(x_1 + x_2) - v_1(x_1))$  and, thus, the final price of item 2 is  $R_2$ . More formally, we have  $\forall p \in I_\alpha^1, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta], f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = 0$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = R_2$  with  $R_2 = \min(v_2(x_2), v_1(x_1 + x_2) - v_1(x_1))$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in [v_1(x_2); v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2))]$ ,

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_1)^2 + (R_2 - p_2)^2 \\ &= (v_1(x_1 + x_2) - \alpha)^2 + (R_2 - \alpha)^2 \end{aligned} \quad (\text{B.14})$$

Thus,  $\forall p \in I_\alpha^1, g_\Gamma(p) \geq (v_2(x_1 + x_2) - v_2(x_2))^2 > 0$ .

- Suppose  $p \in I_\alpha^2$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_1(x_1)$ . As  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1)$ , player 2 will always bid on both items if their price is below  $v_1(x_1 + x_2) - v_1(x_1)$ . As  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 also always bids on both items. Thus, the price

of each item rises symmetrically. As  $p_2 + \eta_1 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 - \eta_1 > v_1(x_1 + x_2) - v_1(x_1)$ , the complementarity surplus of player 1 will be entirely consumed when the price of item 2 reaches  $p_2 + \eta_1$ . Player 1 will then drop out of the auction and the final price of each item will be  $p_2 + \eta_1$ . More formally,  $\forall p \in I_\alpha^2$ ,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in [A; v_1(x_1 + x_2) - v_1(x_1))$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \quad (\text{B.15})$$

Thus,  $\forall p \in I_\alpha^2$ ,  $g_\Gamma(p) > 0$ .

- Suppose  $p \in I_\alpha^3$  and, thus,  $p = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ . As  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1)$ , player 2 will always bid on both items if their price is below  $v_1(x_1 + x_2) - v_1(x_1)$ . As  $p_1 - \eta_1 < v_1(x_1)$  and  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1) < \frac{v_1(x_1 + x_2)}{2} < v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus has been entirely consumed by the rise in price of item 2. This happens when the price of item 2 reaches  $p_2$ . Thus, player 2 acquires item 2 at price  $p_2$ . Then, player 1 will only bid on item 1 until its price reaches her stand-alone value  $p_1 = v_1(x_1)$ . As  $v_1(x_1) \leq v_2(x_1)$ , player 2 will also continue to bid on item 1 and, then, acquire it for a final price of  $p_1$ . More formally, for  $p = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ ,  $\exists \eta > 0$  such that  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = p_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2$ . Thus, we have:

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = 0 \quad (\text{B.16})$$

Thus,  $g_\Gamma(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) = 0$ .

We have shown that:

- If  $I_\alpha^1 \neq \emptyset$ ,  $g_\Gamma|_{O_\Gamma^{min}}$  is lower-bounded by a strictly positive term on  $I_\alpha^1$ .

- $\forall p \in I_\alpha^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [A, v_1(x_1 + x_2) - v_1(x_1)]\}, g_\Gamma(p) > 0.$
- $g_\Gamma(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) = 0$

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))\}.$

□

**Proof. Sequence convergence:**

Let's first show that:

$$\begin{aligned} p &\in [0, v_1(x_1)) \times [0, v_1(x_1 + x_2) - v_1(x_1)) \\ \implies f_\Gamma(p_1, p_2) &= (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) \end{aligned} \quad (\text{B.17})$$

Suppose  $p \in [0, v_1(x_1)) \times [0, v_1(x_1 + x_2) - v_1(x_1))$ . As  $p_1 < v_2(x_1)$ ,  $p_2 < v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_1)$  and  $v_1(x_1 + x_2) - v_1(x_1) < v_2(x_1)$ , player 2 will bid on both items if their price is below  $v_1(x_1 + x_2) - v_1(x_1)$ . As  $p_1 < v_1(x_1)$  and  $p_2 < v_1(x_1 + x_2) - v_1(x_1) < v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed by item 2. Thus, the price of each item rises symmetrically. Player 1 will stop bidding on item 2 when its price reaches  $v_1(x_1 + x_2) - v_1(x_1)$  and will continue to bid on item 1 until its price reaches  $v_1(x_1)$ . Player 2 will then acquire item 2 for  $v_1(x_1 + x_2) - v_1(x_1)$  and will continue bidding on item 1 until its price reaches  $v_2(x_1 + x_2) - v_2(x_2)$ . As  $v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$ , player 2 acquires item 1 for  $v_1(x_1)$ . Thus,  $f_\Gamma(p_1, p_2) = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ .

As  $P_0 = (0, 0) < (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$  component-wise and  $P_{t+1} = \frac{1}{t+1} f_\Gamma(P_t) + (1 - \frac{1}{t+1}) P_t$ , by a simple induction we have that  $\forall t \in \mathbb{R}_+, P_t \in [0, v_1(x_1)) \times [0, v_1(x_1 + x_2) - v_1(x_1))$  and, thus,  $\forall t \in \mathbb{R}_+, f_\Gamma(P_t) = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ .

The above results directly imply that  $P_t \xrightarrow[t \rightarrow +\infty]{} (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

ii. **Second sub-case:**

$$* v_2(x_1 + x_2) - v_2(x_1) < v_1(x_1 + x_2) - v_1(x_1) \text{ (or respectively } v_2(x_1 + x_2) - v_2(x_2) < v_1(x_1 + x_2) - v_1(x_2))$$

*Proof. Frontier prediction of final prices:*

We decompose  $I_\alpha$  in 4 disjoint sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2); A]\}$
- $I_\alpha^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [A, v_2(x_1 + x_2) - v_2(x_1)]\}$
- $I_\alpha^3 = \{(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))\}$
- $I_\alpha^4 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$

with  $A = \max(v_1(x_2), v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)))$  such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3 \cup I_\alpha^4$ . We will calculate the infimum of  $g_\Gamma$  on each of these four sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- If  $v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)) \leq v_1(x_2)$ , i.e.,  $A = v_1(x_2)$ , then  $I_\alpha^1 = \emptyset$ . Otherwise, we have  $v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)) > v_1(x_2)$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p \in I_\alpha^1$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_2)$  and, thus, player 2 won't bid on item 1. The final price of item 1 is then 0. However, player 1 predicts that item 1 will be sold for  $p_1 - \eta_1 > v_1(x_1)$  and, hence, she thinks that she is exposed of  $p_1 - \eta_1 - v_1(x_1)$ . It is then beneficial for player 1 to acquire item 2 if its price is below  $(v_1(x_1 + x_2) - (p_1 - \eta_1)) + (p_1 - \eta_1 - v_1(x_1))$ . The first term comes from the surplus of acquiring item 2 given player 1 has acquired item 1 and the second term comes from its actual predicted exposure. Player 2 will however continue to bid on item 2 until its price reaches  $v_2(x_2)$ . Thus, if player 1 stops bidding on item 2 first, then its final price is  $v_1(x_1 + x_2) - v_1(x_1)$ . If player 2 stops bidding on item 2 first, then its final price is  $v_2(x_2)$ . We set  $R_2 = \min(v_2(x_2), v_1(x_1 + x_2) - v_1(x_1))$  and, thus,



the final price of item 2 is  $R_2$ . More formally, we have  $\forall p \in I_\alpha^1, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta], f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = 0$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = R_2$  with  $R_2 = \min(v_2(x_2), v_1(x_1 + x_2) - v_1(x_1))$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in [v_1(x_2); v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2))]$ , we have

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_1)^2 + (R_2 - p_2)^2 \\ &= (v_1(x_1 + x_2) - \alpha)^2 + (R_2 - \alpha)^2 \end{aligned} \quad (\text{B.18})$$

The above function is minimised for  $\alpha = \frac{R_2 + v_1(x_1 + x_2)}{2} \geq \frac{v_1(x_1 + x_2)}{2} > v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2))$ . Thus,  $\forall [v_1(x_2), v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2))]$  with  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$ , a lower bound of the above function is:

$$\begin{aligned} &(v_2(x_1 + x_2) - v_2(x_2))^2 + (v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_2)) - R_2)^2 \\ &\geq (v_2(x_1 + x_2) - v_2(x_2))^2 \\ &\geq \left(\frac{v_1(x_1 + x_2)}{2}\right)^2 \end{aligned} \quad (\text{B.19})$$

$$\forall p \in I_\alpha^1, g_\Gamma(p) \geq \left(\frac{v_1(x_1 + x_2)}{2}\right)^2$$

- Suppose  $p \in I_\alpha^2$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_2 + \eta_1 < v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_2(x_1 + x_2) - v_2(x_1) < v_2(x_1)$ , player 2 will always bid on both items if their price is below  $v_2(x_1 + x_2) - v_2(x_1)$ . As  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 also always bids on both items. Thus, the price of each item rises symmetrically. As  $p_2 + \eta_1 < \frac{v_1(x_1 + x_2)}{2}$  and  $p_1 - \eta_1 > \frac{v_1(x_1 + x_2)}{2}$ , the complementarity surplus of player 1 will be entirely consumed when the price of item 2 reaches  $p_2 + \eta_1$ . Player 1 will then drop out of the auction and the final price of each item will be  $p_2 + \eta_1$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta], f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ .

Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in [A, v_2(x_1 + x_2) - v_2(x_1)]$ , we have

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \quad (\text{B.20})$$

The function  $\alpha \rightarrow \lim_{\eta \rightarrow 0^+} \frac{1}{\eta^2} \int_{\Omega=[0,\eta]^2} (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 d\eta_1 d\eta_2$  admits a minimum for  $\alpha = \frac{v_1(x_1 + x_2)}{2} \geq v_2(x_1 + x_2) - v_2(x_1)$ .

Thus, the only possible candidate given by the study of  $I_\alpha^2$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .

- Suppose  $p \in I_\alpha^3$  and, thus,  $p = (v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_2(x_1 + x_2) - v_2(x_1) < v_2(x_1)$ , player 2 will always bid on both items if their price is below  $v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_2(x_1 + x_2) - v_2(x_1) < v_1(x_1 + x_2) - v_1(x_2)$ ,  $v_2(x_1 + x_2) - v_2(x_1) < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 also always bids on both items if their price is below  $v_2(x_1 + x_2) - v_2(x_1)$ . Thus, the price of each item rises symmetrically. As  $p_2 = v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1)$ , the complementarity surplus of player 2 will be entirely consumed when the price of item 2 reaches  $p_2$ . Player 1 will then acquire item 2 at price  $p_2$ . Player 1 is then currently exposed of  $p_2 - v_1(x_2)$ . It is then beneficial for player 1 to acquire item 1 if its price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . Player 2 will continue bidding on item 1 if its price is below  $v_2(x_1)$ . Thus, if player 1 stops bidding first on item 1, then its final price is  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding first on item 1, then its final price is  $v_2(x_1)$ . We set  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$  and, thus, the final price of item 1 is  $R_1$ . More formally,  $\exists \eta > 0$  such

$\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_{\Gamma}^2(p_1 - \eta_1, p_2 - \eta_2) = p_2$  and  $f_{\Gamma}^1(p_1 - \eta_1, p_2 - \eta_2) = R_1$ . Thus, we have:

$$\begin{aligned}
& \|f_{\Gamma}(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 \\
&= (R_1 - p_1)^2 \\
&= \min(v_2(x_1 + x_2) - v_1(x_1 + x_2), v_2(x_1 + x_2) - v_1(x_2) - v_2(x_1))^2 \\
&\leq (v_2(x_1 + x_2) - v_1(x_2) - v_2(x_1))^2 \\
&< \left(\frac{v_1(x_1 + x_2)}{2}\right)^2
\end{aligned} \tag{B.21}$$

Thus,  $g_{\Gamma}(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1)) < \left(\frac{v_1(x_1 + x_2)}{2}\right)^2$ .

- Suppose  $p \in I_{\alpha}^4$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > v_2(x_1 + x_2) - v_2(x_1)$  and, thus, player 2 won't bid on item 2. The final price of item 2 is then 0. However, player 1 predicts that item 2 will be sold for  $p_2 - \eta_2 > v_1(x_2)$  and, hence, she thinks that she is exposed of  $p_2 - \eta_2 - v_1(x_2)$ . It is then beneficial for player 1 to acquire item 1 if its price is below  $(v_1(x_1 + x_2) - (p_2 - \eta_2)) + (p_2 - \eta_2 - v_1(x_2))$ . The first term comes from the surplus of acquiring item 1 given player 1 has acquired item 2 and the second term comes from its actual predicted exposure. Player 2 will however continue to bid on item 1 until its price reaches  $v_2(x_1)$ . Thus, if player 1 stops bidding on item 1 first, then its final price is  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding on item 1 first, then its final price is  $v_2(x_1)$ . We set  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$  and, thus, the final price of item 1 is  $R_1$ . More formally, we have  $\forall p \in I_{\alpha}^4$ ,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_{\Gamma}^1(p_1 - \eta_1, p_2 - \eta_2) = R_1$  and  $f_{\Gamma}^2(p_1 - \eta_1, p_2 - \eta_2) = 0$  with  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1))$ , we have:

$$\begin{aligned}
\|f_{\Gamma}(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (R_1 - p_1)^2 + (p_2)^2 \\
&= (R_1 + \alpha - v_1(x_1 + x_2))^2 + (\alpha)^2
\end{aligned} \tag{B.22}$$

The above function is minimised for  $\alpha = \frac{1}{2} \min(v_1(x_2), v_1(x_1 + x_2) - v_2(x_1)) \leq v_2(x_1 + x_2) - v_2(x_1)$ .

Thus, the only possible candidate given by the study of  $I_\alpha^4$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .

We have shown that:

- $g_\Gamma(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1)) < (\frac{v_1(x_1 + x_2)}{2})^2$
- The only possible candidate given by the study of  $I_\alpha^2$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .
- The only possible candidate given by the study of  $I_\alpha^4$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .
- If  $I_\alpha^1 \neq \emptyset$ ,  $\forall p \in I_\alpha^1$ ,  $g_\Gamma(p) \geq (\frac{v_1(x_1 + x_2)}{2})^2$

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))\}$ .

□

**Proof. Sequence convergence:**

Let's set  $V_1 = v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1))$ ,  $V_2 = v_2(x_1 + x_2) - v_2(x_1)$  and  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ .

Let's define:

- $S_1 = \{p \in \mathbb{R}_+^2, p_1 \leq V_1, p_2 < V_2\}$
- $S_2 = \{p \in \mathbb{R}_+^2, V_1 < p_1 < R_1, p_1 + p_2 < v_1(x_1 + x_2)\}$
- $S_3 = \{p \in \mathbb{R}_+^2, V_1 < p_1 < R_1, p_2 < V_2, p_1 + p_2 \geq v_1(x_1 + x_2)\}$

Let's show the three following points:

- $\forall p \in S_1, f_1^1(p) = R_1 \geq V_1, f_1^2(p) = V_2$
- $\forall p \in S_2, f_1^1(p) = f_1^2(p) = v_1(x_1 + x_2) - p_1 \leq V_2 < V_1$
- $\forall p \in S_3, f_1^1(p) = f_1^2(p) = 0 < V_2 < V_1$

Suppose  $p \in S_1$ . As  $V_2 < V_1 \leq v_2(x_1)$ , player 2 will always bid on both items if their price is below  $V_2$ . As  $V_2 < v_1(x_1 + x_2) - v_1(x_1)$ ,  $V_2 < V_1 < v_1(x_1 + x_2) - v_1(x_2)$  and  $V_2 < \frac{v_1(x_1 + x_2)}{2}$ , player 1 also always bids on both items if their price is below  $V_2$ . Thus, the price of each item rises symmetrically. As  $V_2 < v_2(x_1)$ , the complementarity surplus of player 2 will be entirely consumed when the price of item 2 reaches  $V_2$ . Player 1 will then acquire item 2 at price  $V_2$ . Player 1 is then currently exposed of  $V_2 - v_1(x_2)$ . It is then beneficial for player 1 to acquire item 1 if its price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . Player 2 will continue bidding on item 1 if its price is below  $v_2(x_1)$ . Thus, if player 1 stops bidding first on item 1, then its final price is  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding first on item 1, then its final price is  $v_2(x_1)$ . The final price of item 1 is then  $R_1$ . Thus, we have  $f_1^1(p) = R_1 \geq V_1, f_1^2(p) = V_2$ .

This proves the first point.

Suppose  $p \in S_2$ . As  $p_2 < V_2, p_1 < v_2(x_1 + x_2) - v_2(x_2)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items until her complementarity surplus is consumed entirely. As  $p_2 < v_1(x_1 + x_2) - v_1(x_1), p_1 < v_1(x_1 + x_2) - v_1(x_2)$  and  $p_1 + p_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items until her complementarity surplus is consumed entirely. As  $p_2 < v_1(x_1 + x_2) - p_1 < V_2 < V_1 < p_1$ , player 1 complementarity surplus will be entirely consumed when the price of item 2 reaches  $v_1(x_1 + x_2) - p_1$ . Player 1 will then drop out of the auction and player 2 will acquire each item for  $v_1(x_1 + x_2) - p_1$ . Thus, we have  $f_1^1(p) = f_1^2(p) = v_1(x_1 + x_2) - p_1 < V_2 < V_1$ .

This proves the second point.

Suppose  $p \in S_3$ . As  $p_1 + p_2 \geq v_1(x_1 + x_2), v_1(x_1) < V_1 < p_1$  and  $v_1(x_2) \leq v_1(x_1 + x_2) - R_1 < p_2$ , player 1 won't bid on any items. Thus, player 2 will acquire both items for zero. Thus,  $f_1^1(p) = f_1^2(p) = 0 < V_2 < V_1$ .

This proves the third point.

Regarding sequence  $P_{t+1} = \frac{1}{t+1}f_{\Gamma}(P_t) + (1 - \frac{1}{t+1})P_t$ , the three points above directly imply that:

- If  $P_t \in S_1$ , then, as  $f_{\Gamma}^1(P_t) = R_1$  and  $f_{\Gamma}^2(P_t) = V_2$ ,  $P_{t+1} \in S_1 \cup S_2 \cup S_3$ .
- If  $P_t \in S_2$ , then, as  $f_{\Gamma}^1(P_t) = f_{\Gamma}^2(P_t) = v_1(x_1 + x_2) - p_1 \leq V_2 < V_1$ ,  $P_{t+1} \in S_1 \cup S_2$ .
- If  $P_t \in S_3$ , then, as  $f_{\Gamma}^1(p) = f_{\Gamma}^2(p) = 0$ ,  $P_{t+1} \in S_1 \cup S_2 \cup S_3$ .

As  $P_0 = (0, 0) \in S_1$ , we have by induction that  $\forall t \in \mathbb{R}_+$ ,  $P_t \in S_1 \cup S_2 \cup S_3$ .

Moreover,

$$\begin{cases} \text{If } P_t^1 \leq V_1, P_t^1 \in S_1 \text{ and, thus, } f_{\Gamma}^1(P_t) > V_1 \\ \text{If } P_t^1 > V_1, P_t^1 \in S_2 \cup S_3 \text{ and, thus, } f_{\Gamma}^1(P_t) < V_1 \end{cases} \quad (\text{B.23})$$

This directly implies that  $P_t^1 \xrightarrow[t \rightarrow +\infty]{} V_1$ .

Let's show that  $P_t^2 \xrightarrow[t \rightarrow +\infty]{} V_2$ , i.e.,  $\forall \eta > 0$ ,  $\exists t_0 \in \mathbb{R}_+$ ,  $\forall t \geq t_0$ ,  $P_t^2 \in [V_2 - \eta, V_2)$ .

Let's fix  $\eta > 0$ .

Let's fix  $t_1 = \lceil 2(\frac{V_2}{\eta} - 1) \rceil$ .

As  $P_t^1 \xrightarrow[t \rightarrow +\infty]{} V_1$ ,  $\exists t_2 \in \mathbb{R}_+$ ,  $\forall t \geq t_2$ ,  $P_t^1 \in [V_1 - \frac{\eta}{2}, V_1 + \frac{\eta}{2}]$ . Let's fix such a  $t_2$ .

Suppose  $t \geq \max(t_1, t_2)$ . If  $P_t^2 < V_2 - \frac{\eta}{2}$ , then  $P_t \in S_1 \cup S_2$  and, thus,  $f_{\Gamma}^2(P_t) \geq V_2 - \frac{\eta}{2}$ . Hence,  $\exists t_3 > \max(t_1, t_2)$ ,  $P_t^2 \in [V_2 - \eta, V_2)$ .

Let's fix such a  $t_3$ .

We will show by induction that  $\forall t \geq t_3$ ,  $P_t^2 \in [V_2 - \eta, V_2)$  and, hence,  $P_t^2 \xrightarrow[t \rightarrow +\infty]{} V_2$ .

Suppose  $t > t_3$  and  $P_t^2 \in [V_2 - \eta, V_2)$ . Three cases can occur.

- If  $P_t \in S_1$ , then  $f_\Gamma^2(P_t) = V_2$  and, thus,  $P_{t+1}^2 \in [V_2 - \eta, V_2)$ .
- If  $P_t \in S_2$ , then  $f_\Gamma^2(P_t) \geq V_2 - \frac{\eta}{2}$  and, thus,  $P_{t+1}^2 \in [V_2 - \eta, V_2)$ .
- If  $P_t \in S_3$ , then  $P_t^2 \geq V_2 - \frac{\eta}{2}$  and  $f_\Gamma^2(P_t) = 0$ . We have:

$$\begin{aligned}
 P_{t+1}^2 &= \frac{t}{t+1} P_t^2 \geq \frac{t}{t+1} (V_2 - \frac{\eta}{2}) \\
 &\geq \frac{2(\frac{V_2}{\eta} - 1)}{2(\frac{V_2}{\eta} - 1) + 1} (V_2 - \frac{\eta}{2}) \quad (\text{as } t > t_3 > t_1 = \lceil 2(\frac{V_2}{\eta} - 1) \rceil) \\
 &\geq \frac{(\frac{V_2}{\eta} - 1)}{\frac{1}{\eta}(2V_2 - \eta)} (2V_2 - \eta) \\
 &\geq V_2 - \eta
 \end{aligned}
 \tag{B.24}$$

Thus,  $P_{t+1}^2 \in [V_2 - \eta, V_2)$ .

We have shown by induction that  $\forall t \geq t_3, P_t^2 \in [V_2 - \eta, V_2)$ . Hence,  $P_t^2 \xrightarrow[t \rightarrow +\infty]{} V_2$ .

We have shown that  $P_t \xrightarrow[t \rightarrow +\infty]{} (V_1, V_2)$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

## 2. Second disposition:

$$\left\{ \begin{array}{l} v_1(x_1) \leq v_2(x_1) \\ v_1(x_2) \geq v_2(x_2) \\ v_1(x_1 + x_2) \leq v_2(x_1 + x_2) \end{array} \right. \tag{B.25}$$

### a) First case:

$$* \quad v_2(x_1) \geq v_1(x_1 + x_2) - v_1(x_2)$$

$$* v_1(x_2) \geq v_2(x_1 + x_2) - v_2(x_1)$$

**Proof. Frontier prediction of final prices:**

In this proof, we are going to calculate the frontier prediction of final prices and, thus, calculate the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_\Gamma|_{O_\Gamma^{min}}$  for any  $\gamma > 0$ . Lets define:

- $O_\Gamma^1 = \{(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1))\}$
- $O_\Gamma^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1)]\}$
- $O_\Gamma^3 = \{(v_2(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_2(x_2), v_2(x_1 + x_2) - (v_1(x_1 + x_2) - v_1(x_2))]\}$

Lets show that:

- If  $v_1(x_2) > v_2(x_1 + x_2) - v_2(x_1)$ , then  $O_\Gamma^{min} = O_\Gamma^1$ .
- If  $v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$  and  $v_2(x_1) > v_1(x_1 + x_2) - v_1(x_2)$ , then  $O_\Gamma^{min} = O_\Gamma^1 \cup O_\Gamma^2$ .
- If  $v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$  and  $v_2(x_1) = v_1(x_1 + x_2) - v_1(x_2)$ , then  $O_\Gamma^{min} = O_\Gamma^1 \cup O_\Gamma^2 \cup O_\Gamma^3$ .

This will be shown in 3 steps:

- $O_\Gamma^1 \subset O_\Gamma$ ,  $O_\Gamma^2 \subset O_\Gamma$  and  $O_\Gamma^3 \subset O_\Gamma$
- $\forall p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 < v_1(x_1 + x_2) - v_1(x_2) + v_2(x_1 + x_2) - v_2(x_1)$ , then  $p \notin O_\Gamma$ .
- $\forall p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 = v_1(x_1 + x_2)$ , we will see at which conditions  $p \in O_\Gamma$ .

If  $p \in O_\Gamma^1$ , player 1 won't bid on item 1 and player 2 won't bid on item 2 as  $p_1 \geq v_1(x_1 + x_2) - v_1(x_2)$  and  $p_2 \geq v_2(x_1 + x_2) - v_2(x_1)$ . Thus,  $f_\Gamma^1(p) = f_\Gamma^2(p) = 0$  and  $p \in O_\Gamma$ .



If  $p \in O_{\Gamma}^2$ , player 2 won't bid on item 2 as  $p_2 \geq v_2(x_1 + x_2) - v_2(x_1)$ . Thus,  $f_{\Gamma}^2(p) = 0$ . As  $p_1 + p_2 \geq v_1(x_1 + x_2)$  and  $p_1 \geq v_1(x_1)$ , player 1 won't bid on item 1. Thus,  $f_{\Gamma}^1(p) = 0$  and  $p \in O_{\Gamma}$ .

If  $p \in O_{\Gamma}^3$ , player 1 won't bid on item 1 as  $p_1 \geq v_1(x_1 + x_2) - v_1(x_2)$  and player 2 won't bid on item 2 as  $p_1 + p_2 \geq v_2(x_1 + x_2)$  and  $p_2 \geq v_2(x_2)$ . Thus,  $f_{\Gamma}^1(p) = f_{\Gamma}^2(p) = 0$  and  $p \in O_{\Gamma}$ .

The first step is verified.

Suppose  $p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 < v_1(x_1 + x_2) - v_1(x_2) + v_2(x_1 + x_2) - v_2(x_1) \leq v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ . Then, we have:

- If  $p_1 \in [0, v_1(x_1))$ , then, regardless of  $p_2$ ,  $f_{\Gamma}^1(p) \geq v_1(x_1)$  as  $p_1 < v_1(x_1) \leq v_2(x_1)$  and, thus,  $p \notin O_{\Gamma}$ .
- If  $p_1 \in [v_1(x_1), v_1(x_1 + x_2) - v_1(x_2))$ , as  $p_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 < v_1(x_2)$  and  $p_1 + p_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed by item 1 as player 2 will never play above  $v_2(x_1 + x_2) - v_2(x_1) \leq v_1(x_2)$  on item 2. As  $p_1 < v_2(x_1)$ , player 2 will always bid on item 1 until its price reaches  $v_2(x_1)$  regardless of  $p_2$ . Thus,  $f_{\Gamma}^1(p) = v_1(x_1 + x_2) - v_1(x_2)$  and  $p \notin O_{\Gamma}$ .
- If  $p_1 \in [v_1(x_1 + x_2) - v_1(x_2), v_2(x_1))$ , then player 1 won't bid on item 1 as  $p_1 \geq v_1(x_1 + x_2) - v_1(x_2)$  and player 2 will acquire it without consuming any of her complementarity surplus as  $p_1 < v_2(x_1)$ . As  $p_2 < v_2(x_1 + x_2) - v_2(x_1)$ , player 2 will bid on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_2(x_1 + x_2) - v_2(x_1) \leq v_1(x_2)$ ,  $f_{\Gamma}^2(p) = v_2(x_1 + x_2) - v_2(x_1) > p_2$  and, thus,  $p \notin O_{\Gamma}$ .
- If  $p_1 \in [v_2(x_1), A)$  with  $A = \min(v_1(x_1 + x_2) - v_1(x_2) + v_2(x_1 + x_2) - v_2(x_1), v_2(x_1 + x_2) - v_2(x_2))$ , then player 2 will acquire item 1 for a price of 0. Thinking that she is exposed if she doesn't obtain both items, player 2 will bid up to  $v_2(x_1 + x_2) - v_2(x_1)$  on item 2. As  $p_2 < v_1(x_2)$ , player 1 will bid up to  $v_1(x_2)$  on item 2. Thus,  $f_{\Gamma}^2(p) = v_2(x_1 + x_2) - v_2(x_1) > p_2$  and, thus,  $p \notin O_{\Gamma}$ .

- If  $v_2(x_1 + x_2) - v_2(x_2) < v_1(x_1 + x_2) - v_1(x_2) + v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 \in [v_2(x_1 + x_2) - v_2(x_2), v_1(x_1 + x_2) - v_1(x_2) + v_2(x_1 + x_2) - v_2(x_1)]$ , then, regardless of  $p_1$ ,  $f_\Gamma^2(p) \geq v_2(x_2)$  as  $p_2 < v_2(x_2) \leq v_1(x_2)$  and, thus,  $p \notin O_\Gamma$ .

The second step is verified.

Suppose  $p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 = v_1(x_1 + x_2) - v_1(x_2) + v_2(x_1 + x_2) - v_2(x_1) \leq v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ . Then, we have:

- If  $p_2 \in [0, v_2(x_2))$ , then, regardless of  $p_1$ ,  $f_\Gamma^2(p) \geq v_2(x_2)$  as  $p_2 < v_2(x_2) \leq v_1(x_2)$  and, thus,  $p \notin O_\Gamma$ .
- If  $p_2 \in [v_2(x_2), v_2(x_1 + x_2) - v_2(x_1))$ , then we need to study two different cases:
  - If  $v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$  and  $v_2(x_1) = v_1(x_1 + x_2) - v_1(x_2)$ , as  $p_1 + p_2 = v_2(x_1 + x_2)$ ,  $p_2 \geq v_2(x_2)$  and  $p_1 \geq v_1(x_1 + x_2) - v_1(x_2)$ , we have  $p \in O_\Gamma^3 \subset O_\Gamma$ .
  - Otherwise, as  $p_1 \geq v_1(x_1 + x_2) - v_1(x_2)$ , player 1 won't bid on item 1. As  $p_1 + p_2 < v_2(x_1 + x_2)$ ,  $p_2 < v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 < v_2(x_1 + x_2) - v_2(x_2)$ , player 2 will bid on both items until her complementarity surplus is entirely consumed. As player 2 acquires item 1 at a price of zero, she will bid on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1)$ . As  $p_2 < v_1(x_2)$ , player 1 will bid on item 2 if its price is below  $v_1(x_2)$ . Thus,  $f_\Gamma^2(p) = v_2(x_1 + x_2) - v_2(x_1) > p_2$  and  $p \notin O_\Gamma$ .
- If  $p_2 = v_2(x_1 + x_2) - v_2(x_1)$ , then  $p \in O_\Gamma^1 \subset O_\Gamma$ .
- If  $p_2 \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1)]$ , then we need to study two different cases:
  - If  $v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$ , as  $p_1 + p_2 = v_1(x_1 + x_2)$ ,  $p_2 \geq v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 \geq v_1(x_1)$ , we have  $p \in O_\Gamma^2 \subset O_\Gamma$ .
  - Otherwise, as  $p_2 > v_2(x_1 + x_2) - v_2(x_1)$ , player 2 won't bid on item 2. As  $p_1 + p_2 < v_1(x_1 + x_2)$ ,  $p_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 < v_1(x_1 + x_2) - v_1(x_2)$ , player 1 will bid on both items until her complementarity

surplus is entirely consumed. As player 1 acquires item 2 at a price of zero, she will bid on item 1 until its price reaches  $v_1(x_1 + x_2) - v_1(x_2)$ . As  $p_1 < v_2(x_1)$ , player 2 will bid on item 1 if its price is below  $v_2(x_1)$ . Thus,  $f_\Gamma^1(p) = v_1(x_1 + x_2) - v_1(x_2) > p_1$  and  $p \notin O_\Gamma$ .

- If  $p_2 \in (v_1(x_1 + x_2) - v_1(x_1), v_1(x_1 + x_2) - v_1(x_2) + v_2(x_1 + x_2) - v_2(x_1))$ , then, as  $p_1 < v_1(x_1)$ , we have  $f_\Gamma^1(p) \geq v_1(x_1)$  regardless of  $p_2$  and, thus,  $p \notin O_\Gamma$ .

The third step is verified.

We have shown that:

- If  $v_1(x_2) > v_2(x_1 + x_2) - v_2(x_1)$ , then  $O_\Gamma^{min} = O_\Gamma^1$ .
- If  $v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$  and  $v_2(x_1) > v_1(x_1 + x_2) - v_1(x_2)$ , then  $O_\Gamma^{min} = O_\Gamma^1 \cup O_\Gamma^2$ .
- If  $v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$  and  $v_2(x_1) = v_1(x_1 + x_2) - v_1(x_2)$ , then  $O_\Gamma^{min} = O_\Gamma^1 \cup O_\Gamma^2 \cup O_\Gamma^3$ .

Lets determine for the three above cases the frontier prediction of final prices.

- If  $v_1(x_2) > v_2(x_1 + x_2) - v_2(x_1)$ , then  $O_\Gamma^{min} = O_\Gamma^1 = \{(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1))\}$ . Thus,  $O_\Gamma^{min*} = \{(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1))\}$ .
- If  $v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$  and  $v_2(x_1) > v_1(x_1 + x_2) - v_1(x_2)$ , then  $O_\Gamma^{min} = O_\Gamma^1 \cup O_\Gamma^2$ . We can rewrite  $O_\Gamma^{min}$  as  $I_\alpha = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1)]\}$ . From there, we are going to calculate the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_\Gamma|_{O_\Gamma^{min}}$  for any  $\gamma > 0$ . To do so, we need to obtain the closed-form of  $f_\Gamma$  for every slight underestimation of elements in  $I_\alpha$ . We decompose  $I_\alpha$  in 2 disjoints sets

$$- I_\alpha^1 = \{(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1))\}$$

$$- I_\alpha^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$$

such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2$ . We will calculate the infimum of  $g_\Gamma$  on each of these two sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- Suppose  $p \in I_\alpha^1$  and, thus,  $p = (v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1))$ . Lets calculate the final prices of  $\Gamma$  if both players play the point prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ . As  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2) < v_2(x_1)$  and  $p_2 - \eta_2 < v_2(x_1 + x_2) - v_2(x_1) < v_1(x_2)$ , player 1 and player 2 will bid on item 1 until its price reaches  $v_1(x_1 + x_2) - v_1(x_2)$  and on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1)$ . Then, player 1 will stop bidding on the first item and player 2 will stop bidding on the second item. Thus, player 2 acquires item 1 for  $v_1(x_1 + x_2) - v_1(x_2)$  and player 1 acquires item 2 for  $v_2(x_1 + x_2) - v_2(x_1)$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = p_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2$ . Thus, we have

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = 0 \quad (\text{B.26})$$

Thus,  $g_\Gamma(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1)) = 0$ .

- Suppose  $p \in I_\alpha^2$ . Lets calculate the final prices of  $\Gamma$  if both players play the point prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > v_2(x_1 + x_2) - v_2(x_1)$ . As  $p_2 - \eta_2 > v_2(x_1 + x_2) - v_2(x_1)$ , player 2 won't bid on item 2 and will only bid on item 1 if its price is below  $v_2(x_1) \geq v_1(x_1 + x_2) - v_1(x_2)$ . As  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$ ,  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 bids on both items and acquires item 2 for a price of zero. However, player 1 thinks she is exposed of  $p_2 - \eta_2 - v_1(x_2) > 0$ . Thus, player 1 will bid on item 1 until its price reaches  $v_1(x_1 + x_2) - v_1(x_2)$  and, then, stops bidding. Player 2 acquires item 1 for  $v_1(x_1 + x_2) - v_1(x_2)$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = v_1(x_1 + x_2) - v_1(x_2)$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = 0$ . Thus, we have

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = (v_1(x_1 + x_2) - v_1(x_2) - p_1)^2 + (p_2)^2 \quad (\text{B.27})$$

Thus,  $\forall p \in I_\alpha^2, g_\Gamma(p) > 0$ .

We have shown that:

- $g_\Gamma(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1)) = 0$
- $\forall p \in I_\alpha^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1))\}, g_\Gamma(p) > 0$

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1))\}$ .

- If  $v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$  and  $v_2(x_1) = v_1(x_1 + x_2) - v_1(x_2)$ , then  $O_\Gamma^{min} = O_\Gamma^1 \cup O_\Gamma^2 \cup O_\Gamma^3$ . We can rewrite  $O_\Gamma^{min}$  as  $I_\alpha = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_2(x_2), v_1(x_1 + x_2) - v_1(x_1)]\}$ . From there, we are going to calculate the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_\Gamma|_{O_\Gamma^{min}}$  for any  $\gamma > 0$ . To do so, we need to obtain the closed-form of  $f_\Gamma$  for every slight underestimation of elements in  $I_\alpha$ . We decompose  $I_\alpha$  in 3 disjoint sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_2(x_2), v_2(x_1 + x_2) - v_2(x_1)]\}$
- $I_\alpha^2 = \{(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1))\}$
- $I_\alpha^3 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$

such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3$ . We will calculate the infimum of  $g_\Gamma$  on each of these three sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- Suppose  $p \in I_\alpha^1$ . Lets calculate the final prices of  $\Gamma$  if both players play the point prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1)$ . As  $p_1 - \eta_1 > v_1(x_1 + x_2) - v_1(x_2)$ , player 1 won't bid on item 1. As  $p_1 - \eta_1 < v_2(x_1 + x_2) - v_2(x_2)$ ,  $p_2 - \eta_2 < v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items and acquire item 1 for a price of zero. However, player 2 thinks she is exposed of  $p_1 - \eta_1 - v_2(x_1) > 0$  if she doesn't acquire item 2.

Thus, player 2 will bid on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1)$ . Player 1 will bid also on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1)$  as  $p_2 - \eta_2 < v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$ . Thus, the final price of item 2 is  $v_2(x_1 + x_2) - v_2(x_1)$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = 0$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = v_2(x_1 + x_2) - v_2(x_1)$ . Thus, we have:

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = (p_1)^2 + (v_2(x_1 + x_2) - v_2(x_1) - p_2)^2 \quad (\text{B.28})$$

Thus,  $\forall p \in I_\alpha^1, g_\Gamma(p) > 0$

- Suppose  $p \in I_\alpha^2$ . Then, according to the proof in the second case,  $g_\Gamma(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1)) = 0$ .
- Suppose  $p \in I_\alpha^3$ . Then, according to the proof in the second case,  $g_\Gamma(p) > 0$ .

We have shown that:

- $\forall p \in I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_2(x_2), v_2(x_1 + x_2) - v_2(x_1)]\}, g_\Gamma(p) > 0$
- $g_\Gamma(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1)) = 0$
- $\forall p \in I_\alpha^3 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1))\}, g_\Gamma(p) > 0$

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1))\}$ .

We have shown for all cases that  $O_\Gamma^{min*} = \{(v_1(x_1 + x_2) - v_1(x_2), v_2(x_1 + x_2) - v_2(x_1))\}$ .

□

*Proof.* **Sequence convergence:**

Let's set  $V_1 = v_1(x_1 + x_2) - v_1(x_2)$  and  $V_2 = v_2(x_1 + x_2) - v_2(x_1)$ .

Let's first show that:

$$p \in [0, V_1] \times [0, V_2] \implies f_\Gamma(p) = (V_1, V_2) \quad (\text{B.29})$$

Suppose  $p \in [0, V_1] \times [0, V_2]$ . Without any loss of generality, let's suppose that  $V_1 \leq V_2$ . As  $p_1 < V_1 \leq v_2(x_1)$ ,  $p_2 < V_2$  and  $V_1 \leq V_2$ , player 2 will bid on both items if their price is below  $V_1$ . As  $p_1 < V_1$ ,  $p_2 < V_2 \leq v_1(x_2)$  and  $V_1 \leq V_2$ , player 1 will bid on both items until her complementarity surplus has been entirely consumed by item 1. Thus, the price of each item rises symmetrically. Player 1 stops bidding on item 1 when its price reaches  $V_1$  and will continue to bid on item 2 until its price reaches  $v_1(x_2)$ . Thus, player 2 acquires item 1 for  $V_1$  and will continue to bid on item 2 until its price reaches  $V_2$ . As  $v_1(x_2) \geq V_2$ , item 2 will be acquired by player 1 for  $v_1(x_2)$ . Thus,  $f_\Gamma(p) = (V_1, V_2)$ .

As  $P_0 = (0, 0) < (V_1, V_2)$  component-wise and  $P_{t+1} = \frac{1}{t+1} f_\Gamma(P_t) + (1 - \frac{1}{t+1}) P_t$ , by a simple induction we have that  $\forall t \in \mathbb{R}_+$ ,  $P_t \in [0, V_1] \times [0, V_2]$  and, thus,  $\forall t \in \mathbb{R}_+$ ,  $f_\Gamma(P_t) = (V_1, V_2)$ .

The above results directly imply that  $P_t \xrightarrow[t \rightarrow +\infty]{} (V_1, V_2)$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

b) **Second case:**

$$* \quad v_2(x_1 + x_2) - v_2(x_1) \geq v_1(x_2)$$

*Proof.* **Set of smallest overestimated final prices:**

The following result is shared amongst all the sub-cases of the second case of the second disposition. In this proof, we are going to calculate the set of smallest overestimated final prices  $O_\Gamma^{min}$ .

Lets set  $O_\Gamma^1 = \{p, v_1(x_1 + x_2) = p_1 + p_2, v_1(x_1) \leq p_1, v_1(x_2) \leq p_2\}$ . Lets show that  $O_\Gamma^{min} = O_\Gamma^1$  in three steps:

- $O_{\Gamma}^1 \subset O_{\Gamma}$
- $\forall p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 < v_1(x_1 + x_2)$ , then  $p \notin O_{\Gamma}$ . Thus,  $O_{\Gamma}^1 \subset O_{\Gamma}^{min}$ .
- $\forall p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 = v_1(x_1 + x_2)$ , if  $p_1 < v_1(x_1)$  or  $p_2 < v_1(x_2)$  then  $p \notin O_{\Gamma}$ . Thus,  $O_{\Gamma}^1 = O_{\Gamma}^{min}$ .

If  $p \in O_{\Gamma}^1$ , then  $f_{\Gamma}^1(p) = f_{\Gamma}^2(p) = 0$  as player 1 predicts that she will get a negative payoff if she acquires any item. Thus,  $p \in O_{\Gamma}$ . This verifies the first step.

Suppose  $p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 < v_1(x_1 + x_2)$ . Then, we have:

- If  $p_1 \in [0, v_1(x_1))$ , then, regardless of price  $p_2$ ,  $f_{\Gamma}^1(p) \geq v_1(x_1)$  as both players has their stand-alone value for item 1 greater or equal to  $v_1(x_1)$ . Thus,  $p \notin O_{\Gamma}$ .
- If  $p_1 \in [v_1(x_1), v_1(x_1 + x_2) - v_1(x_2))$ , as  $p_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed. Two different scenarios can occur.
  - If  $v_2(x_1 + x_2) - v_2(x_1) > p_2$ , then, as  $p_1 < v_2(x_1 + x_2) - v_2(x_2)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items until her complementarity surplus is consumed. As  $v_1(x_1) < p_1 < v_1(x_1 + x_2) - v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $v_1(x_2) < p_2 < v_2(x_1 + x_2) - v_2(x_1) \leq v_1(x_1 + x_2) - v_1(x_1)$ , player 1 will either entirely consume her complementarity surplus without acquiring any items or player 1 will acquire item 2 for  $v_2(x_1 + x_2) - v_2(x_1)$  and will then bid up to  $v_1(x_1 + x_2) - v_1(x_2)$  on item 1 because of exposure. Either way,  $f_{\Gamma}^1(p) + f_{\Gamma}^2(p) \geq v_1(x_1 + x_2) > p_1 + p_2$  and, thus,  $p \notin O_{\Gamma}$ .
  - If  $v_2(x_1 + x_2) - v_2(x_1) \leq p_2$ , player won't bid on item 2. Player 1 will then acquire item 2 for a price of zero. However, thinking that she is exposed of  $p_2 - v_1(x_2) > 0$ , she will bid on item 1 until its price reaches  $v_1(x_1 + x_2) - v_1(x_2) > p_1$ . Player 2 will bid on item 1 until its price reaches  $v_2(x_1) > p_1$ . Thus,  $f_{\Gamma}^1(p) > p_1$  and  $p \notin O_{\Gamma}$ .
- If  $p_1 \in [v_1(x_1 + x_2) - v_1(x_2), A)$  with  $A = \min(v_1(x_1 + x_2), v_2(x_1 + x_2) - v_2(x_2))$ , then player 1 won't bid on item 1 and only bid on item 2 if its price is below



$v_1(x_2) > p_2$ . As  $p_1 < v_2(x_1 + x_2) - v_2(x_2)$ ,  $p_2 < v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , then player 2 will bid on both items and acquire item 1 for a price of zero. Player 2 will then bid on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1)$  whether she thinks she is exposed or not. Thus,  $f_{\Gamma}^2(p) > p_2$  and  $p \notin O_{\Gamma}$ .

- If  $p_1 \in [A, v_1(x_1 + x_2))$  with  $A = \min(v_1(x_1 + x_2), v_2(x_1 + x_2) - v_2(x_2))$ , then, regardless of  $p_2$ , as  $p_2 < v_2(x_2)$ , we have  $f_{\Gamma}^2(p) > p_2$  and  $p \notin O_{\Gamma}$ .

This verifies the second step.

Suppose  $p \in \mathbb{R}_{+}^2$  such that  $p_1 + p_2 = v_1(x_1 + x_2)$ . Then, we have:

- If  $p_1 \in [0, v_1(x_1))$ , then, regardless of price  $p_2$ ,  $f_{\Gamma}^1(p) \geq v_1(x_1)$  as both players has their stand-alone value for item 1 greater or equal to  $v_1(x_1)$ . Thus,  $p \notin O_{\Gamma}$ .
- If  $p_1 \in [v_1(x_1), v_1(x_1 + x_2) - v_1(x_2)]$ , then  $p \in O_{\Gamma}^1 \subset O_{\Gamma}$ .
- If  $p_1 \in (v_1(x_1 + x_2) - v_1(x_2), A)$  with  $A = \min(v_1(x_1 + x_2), v_2(x_1 + x_2) - v_2(x_2))$ , then player 1 won't bid on item 1 and will bid on item 2 until its price reaches  $v_1(x_2) > p_2$ . As  $p_1 < v_2(x_1 + x_2) - v_2(x_2)$ ,  $p_2 < v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , then player 2 will bid on both items and acquire item 1 for a price of zero. Player 2 will then bid on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1)$  whether she thinks she is exposed or not. Player 2 will then acquire item 2 for  $v_1(x_2)$  as  $v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_1)$ . Thus,  $f_{\Gamma}^2(p) = v_1(x_2) > p_2$  and  $p \notin O_{\Gamma}$ .
- If  $p_1 \in [A, v_1(x_1 + x_2))$  with  $A = \min(v_1(x_1 + x_2), v_2(x_1 + x_2) - v_2(x_2))$ , then, regardless of  $p_2$ , as  $p_2 < v_2(x_2)$ , we have  $f_{\Gamma}^2(p) > p_2$  and  $p \notin O_{\Gamma}$ .

This verifies the third step.

Thus, we have shown that:

$$O_{\Gamma}^{min} = O_{\Gamma}^1 = \{p, v_1(x_1 + x_2) = p_1 + p_2, v_1(x_1) \leq p_1, v_1(x_2) \leq p_2\} \quad (\text{B.30})$$

□

**i. First sub-case:**

$$* v_1(x_2) > v_1(x_1 + x_2) - v_1(x_2)$$

**Proof. Frontier prediction of final prices:**

In this proof, we are going to calculate the frontier prediction of final prices and, thus, calculate the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_\Gamma|_{O_\Gamma^{min}}$  for any  $\gamma > 0$ .

We can rewrite  $O_\Gamma^{min}$  as  $I_\alpha = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2); v_1(x_1 + x_2) - v_1(x_1)]\}$ . From there, we are going to calculate the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_\Gamma|_{O_\Gamma^{min}}$  for any  $\gamma > 0$ . To do so, we need to obtain the closed-form of  $f_\Gamma$  for every slight underestimation of elements in  $I_\alpha$ . We decompose  $I_\alpha$  in 3 disjoint sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))\}$
- $I_\alpha^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (v_1(x_2), B]\}$
- $I_\alpha^3 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (B, v_1(x_1 + x_2) - v_1(x_1)]\}$

with  $B = \min(v_1(x_1 + x_2) - v_1(x_1), v_2(x_1 + x_2) - v_2(x_1))$  such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3$ . We will calculate the infimum of  $g_\Gamma$  on each of these three sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- Suppose  $p \in I_\alpha^1$  and, thus,  $p = (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ . As  $p_2 - \eta_2 < v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items until her complementarity surplus is entirely consumed. As  $p_2 - \eta_2 < v_1(x_2)$  and  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2) < v_1(x_2)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed by item 1. Thus, player 2 acquires item 1 at  $v_1(x_1 + x_2) - v_1(x_2)$ . Player 1 will

then bid on item 2 until its price reaches  $v_1(x_2)$ . Given that player 2 has acquired item 1, she will then bid on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1) \geq v_1(x_2)$ . Thus, the final price of item 2 is  $v_1(x_2)$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = p_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2$ . Thus, we have:

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = 0 \quad (\text{B.31})$$

Thus,  $g_\Gamma(v_1(x_1 + x_2) - v_1(x_2), v_1(x_2)) = 0$ .

- If  $v_1(x_2) = v_2(x_1 + x_2) - v_2(x_1)$ , then  $I_\alpha^2 = \emptyset$ . Otherwise, we have  $v_1(x_2) < v_2(x_1 + x_2) - v_2(x_1)$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p \in I_\alpha^2$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > v_1(x_2)$ . As  $p_2 - \eta_2 < v_2(x_1 + x_2) - v_2(x_1)$ ,  $p_1 - \eta_1 < v_2(x_1 + x_2) - v_2(x_2)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items until her complementarity surplus is entirely consumed. As  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$ ,  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed. Prices will rise symmetrically. As  $v_1(x_1 + x_2) - v_1(x_2) < p_2 - \eta_2$  and  $v_1(x_1 + x_2) - v_1(x_2) > p_1 + \eta_2$ , the complementarity surplus of player 1 will be entirely consumed when the price of item 1 reaches  $p_1 + \eta_2$ . Player 1 will then drop out of the auction and the final price of each item will be  $p_1 + \eta_2$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_1 + \eta_2$ . Thus,  $\forall p \in I_\alpha^2$  with  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_1 + \eta_2 - p_1)^2 + (p_1 + \eta_2 - p_2)^2 \\ &= (v_1(x_1 + x_2) - 2\alpha + \eta_2)^2 + \eta_2^2 \end{aligned} \quad (\text{B.32})$$

Thus,  $\forall p \in I_\alpha^2, g_\Gamma(p) > 0$ .

- If  $v_2(x_1 + x_2) - v_2(x_1) \geq v_1(x_1 + x_2) - v_1(x_1)$ , i.e.,  $B = v_1(x_1 + x_2) - v_1(x_1)$ , then  $I_\alpha^3 = \emptyset$ . Otherwise, we have  $v_2(x_1 + x_2) - v_2(x_1) < v_1(x_1 + x_2) - v_1(x_1)$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p \in I_\alpha^3$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such

that  $p_2 - \eta_2 > v_2(x_1 + x_2) - v_2(x_1)$ . As  $p_2 - \eta_2 > v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 - \eta_1 < v_2(x_1)$ , player 2 won't bid on item 2 and will bid on item 1 until its price reaches  $v_2(x_1)$ . As  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$ ,  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items and will acquire item 2 for a price of zero. However, player 1 thinks that she is exposed of  $p_2 - \eta_1 - v_1(x_2) > 0$  and will continue to bid on item 1 until its price reaches  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding first on item 1, then its final price will be  $v_2(x_1)$ . If player 1 stops bidding first on item 1, then its final price will  $v_1(x_1 + x_2) - v_1(x_2)$ . The final price of item 1 will be  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = R_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = 0$  with  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ . Thus,  $\forall p \in I_\alpha^3$  with  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (R_1 - p_1)^2 + (p_2)^2 \\ &= (R_1 + \alpha - v_1(x_1 + x_2))^2 + \alpha^2 \end{aligned} \quad (\text{B.33})$$

The above function is minimised for  $\alpha = \frac{1}{2} \min(v_1(x_2), v_1(x_1 + x_2) - v_2(x_1)) \leq v_2(x_1 + x_2) - v_2(x_1)$  and, thus,  $\forall p \in I_\alpha^3$ ,  $g_\Gamma(p) \geq (v_2(x_1 + x_2) - v_2(x_1))^2 > 0$

We have shown that:

- $g_\Gamma(v_1(x_1 + x_2) - v_1(x_2), v_1(x_2)) = 0$
- If  $I_\alpha^2 \neq \emptyset$ ,  $\forall p \in I_\alpha^2 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (v_1(x_2), B]\}$ ,  $g_\Gamma(p) > 0$
- If  $I_\alpha^3 \neq \emptyset$ ,  $\forall p \in I_\alpha^3$ ,  $g_\Gamma(p) \geq (v_2(x_1 + x_2) - v_2(x_1))^2 > 0$

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))\}$ .

□

*Proof. Sequence convergence:*

Let's first show that:

$$\begin{aligned} p &\in [0, v_1(x_1 + x_2) - v_1(x_2)] \times [0, v_1(x_2)] \\ \implies f_\Gamma(p_1, p_2) &= (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2)) \end{aligned} \quad (\text{B.34})$$

Suppose  $p \in [0, v_1(x_1 + x_2) - v_1(x_2)] \times [0, v_1(x_2)]$ . As  $v_1(x_1 + x_2) - v_1(x_2) < v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $v_1(x_1 + x_2) - v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $v_1(x_1 + x_2) - v_1(x_2) \leq \frac{v_2(x_1 + x_2)}{2}$ , player 2 will bid on both items if their price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . As  $p_2 < v_1(x_2)$  and  $p_1 < v_1(x_1 + x_2) - v_1(x_2) < v_1(x_2)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed by item 1. Thus, the price of each item rises symmetrically. Player 1 will stop bidding on item 1 when its price reaches  $v_1(x_1 + x_2) - v_1(x_2)$  and will continue to bid on item 2 until its price reaches  $v_1(x_2)$ . Player 2 will then acquire item 1 for  $v_1(x_1 + x_2) - v_1(x_2)$  and will continue bidding on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_1)$ , player 2 acquires item 2 for  $v_1(x_2)$ . Thus,  $f_\Gamma(p_1, p_2) = (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))$ .

As  $P_0 = (0, 0) < (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))$  component-wise and  $P_{t+1} = \frac{1}{t+1} f_\Gamma(P_t) + (1 - \frac{1}{t+1}) P_t$ , by a simple induction we have that  $\forall t \in \mathbb{R}_+$ ,  $P_t \in [0, v_1(x_1 + x_2) - v_1(x_2)] \times [0, v_1(x_2)]$  and, thus,  $\forall t \in \mathbb{R}_+$ ,  $f_\Gamma(P_t) = (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))$ .

The above results directly imply that  $P_t \xrightarrow[t \rightarrow +\infty]{} (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{\min*}$

□

## ii. Second sub-case:

$$* \quad v_1(x_2) \leq v_1(x_1 + x_2) - v_1(x_2)$$

### A. First sub-sub-case:

$$* \quad v_1(x_1) \leq \frac{v_1(x_1 + x_2)}{2}$$

#### First sub-sub-sub-case:

$$* v_2(x_1 + x_2) - v_2(x_1) \geq \frac{v_1(x_1+x_2)}{2}$$

*Proof. Frontier prediction of final prices:*

We decompose  $I_\alpha$  in 4 disjoint sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2), \frac{v_1(x_1+x_2)}{2}]\}$
- $I_\alpha^2 = \{(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})\}$
- $I_\alpha^3 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (\frac{v_1(x_1+x_2)}{2}, B]\}$
- $I_\alpha^4 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (B, v_1(x_1 + x_2) - v_1(x_1))\}$

with  $B = \min(v_1(x_1 + x_2) - v_1(x_1), v_2(x_1 + x_2) - v_2(x_1))$  such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3 \cup I_\alpha^4$ . We will calculate the infimum of  $g_\Gamma$  on each of these four sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- Suppose  $p \in I_\alpha^1$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > \frac{v_1(x_1+x_2)}{2}$ . As  $v_2(x_1 + x_2) - v_2(x_1) \geq \frac{v_1(x_1+x_2)}{2} > p_2 - \eta_2$ ,  $v_2(x_1 + x_2) - v_2(x_2) > p_1 - \eta_1 > \frac{v_1(x_1+x_2)}{2}$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , player 2 will bid on both items if their price is below  $p_2 + \eta_1$ . As  $v_1(x_1 + x_2) - v_1(x_1) > p_2 - \eta_2$ ,  $v_1(x_1 + x_2) - v_1(x_2) > p_1 - \eta_1$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items. Prices will then rise symmetrically. As  $p_1 - \eta_1 > \frac{v_1(x_1+x_2)}{2} > p_2 + \eta_1$ , player 1 will drop out of the auction when the price of each item reaches  $p_2 + \eta_1$ . Thus, player 2 will acquire each item for  $p_2 + \eta_1$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,

$f_{\Gamma}^1(p_1 - \eta_1, p_2 - \eta_2) = f_{\Gamma}^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ . Thus,  $\forall p \in I_{\alpha}^1$  with  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$ , we have:

$$\begin{aligned} \|f_{\Gamma}(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \quad (\text{B.35})$$

The function  $\alpha \rightarrow \lim_{\eta \rightarrow 0^+} \frac{1}{\eta^2} \int_{\Omega=[0, \eta]^2} (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 d\eta_1 d\eta_2$  admits a minimum for  $\alpha = \frac{v_1(x_1 + x_2)}{2}$ .

Thus, the only possible candidate given by the study of  $I_{\alpha}^1$  as an element of the frontier prediction of final prices would be  $(\frac{v_1(x_1 + x_2)}{2}, \frac{v_1(x_1 + x_2)}{2})$ .

- Suppose  $p \in I_{\alpha}^2$  and, thus,  $p = (\frac{v_1(x_1 + x_2)}{2}, \frac{v_1(x_1 + x_2)}{2})$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ . As  $v_2(x_1 + x_2) - v_2(x_1) \geq \frac{v_1(x_1 + x_2)}{2} > p_2 - \eta_2$ ,  $v_2(x_1 + x_2) - v_2(x_2) \geq \frac{v_1(x_1 + x_2)}{2} > p_1 - \eta_1$  and  $v_2(x_1 + x_2) \geq v_1(x_1 + x_2) > p_1 + p_2 - \eta_1 - \eta_2$ , player 2 will bid on both items if their price is below  $\frac{v_1(x_1 + x_2)}{2}$ . As  $p_1 - \eta_1 < \frac{v_1(x_1 + x_2)}{2} \leq v_1(x_1 + x_2) - v_1(x_2)$  and  $p_2 - \eta_2 < \frac{v_1(x_1 + x_2)}{2} \leq v_1(x_1 + x_2) - v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed. Thus, both prices rise symmetrically and player 1 drops out of the auction when they reach  $\frac{v_1(x_1 + x_2)}{2}$ . Player 2 will acquire each item for  $\frac{v_1(x_1 + x_2)}{2}$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_{\Gamma}^1(p_1 - \eta_1, p_2 - \eta_2) = f_{\Gamma}^2(p_1 - \eta_1, p_2 - \eta_2) = p_1 = p_2$ . Thus, we have:

$$\|f_{\Gamma}(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = 0 \quad (\text{B.36})$$

Thus,  $g_{\Gamma}(\frac{v_1(x_1 + x_2)}{2}, \frac{v_1(x_1 + x_2)}{2}) = 0$ .

- If  $\frac{v_1(x_1 + x_2)}{2} = v_2(x_1 + x_2) - v_2(x_1)$ , then  $I_{\alpha}^3 = \emptyset$ . Otherwise,  $\frac{v_1(x_1 + x_2)}{2} < v_2(x_1 + x_2) - v_2(x_1)$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p \in I_{\alpha}^3$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > \frac{v_1(x_1 + x_2)}{2}$ . As  $v_2(x_1 + x_2) - v_2(x_2) \geq \frac{v_1(x_1 + x_2)}{2} > p_1 - \eta_1$ ,  $v_2(x_1 + x_2) - v_2(x_1) > p_2 - \eta_2 > \frac{v_1(x_1 + x_2)}{2}$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , player 2 will bid on both

items if their price is below  $p_1 + \eta_2$ . As  $v_1(x_1 + x_2) - v_1(x_1) > p_2 - \eta_2$ ,  $v_1(x_1 + x_2) - v_1(x_2) > p_1 - \eta_1$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items. Prices will then rise symmetrically. As  $p_2 - \eta_2 > \frac{v_1(x_1+x_2)}{2} > p_1 + \eta_2$ , player 1 will drop out of the auction when the price of each item reaches  $p_1 + \eta_2$ . Thus, player 2 will acquire each item for  $p_1 + \eta_2$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_1 + \eta_2$ . Thus,  $\forall p \in I_\alpha^3$  with  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_1 + \eta_2 - p_1)^2 + (p_1 + \eta_2 - p_2)^2 \\ &= (v_1(x_1 + x_2) - 2\alpha + \eta_2)^2 + \eta_2^2 \end{aligned} \quad (\text{B.37})$$

The function  $\alpha \rightarrow \lim_{\eta \rightarrow 0^+} \frac{1}{\eta^2} \int_{\Omega=[0, \eta]^2} (v_1(x_1+x_2) - 2\alpha + \eta_2)^2 + \eta_2^2 d\eta_1 d\eta_2$  admits a minimum for  $\alpha = \frac{v_1(x_1+x_2)}{2}$ .

Thus, the only possible candidate given by the study of  $I_\alpha^3$  as an element of the frontier prediction of final prices would be  $(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$ .

- If  $v_1(x_1 + x_2) - v_1(x_1) = v_2(x_1 + x_2) - v_2(x_1)$ , i.e.,  $B = v_1(x_1 + x_2) - v_1(x_1)$ , then  $I_\alpha^4 = \emptyset$ . Otherwise,  $v_1(x_1 + x_2) - v_1(x_1) > v_2(x_1 + x_2) - v_2(x_1)$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p \in I_\alpha^4$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > v_2(x_1 + x_2) - v_2(x_1)$ . Thus, as  $p_1 - \eta_1 < v_2(x_1)$ , player 2 will only bid on item 1 until its price reaches  $v_2(x_1)$ . As  $v_1(x_1 + x_2) - v_1(x_1) > p_2 - \eta_2$ ,  $v_1(x_1 + x_2) - v_1(x_2) > p_1 - \eta_1$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items. Player 1 will acquire item 2 for a price of zero. Given that player 1 has acquired item 2, she will bid on item 1 until its price has reached  $v_1(x_1 + x_2) - v_1(x_2)$ . Thus, if player 2 stops bidding first, the final price of item 1 is  $v_2(x_1)$ . If player 1 stops bidding first, the final price of item 1 is  $v_1(x_1 + x_2) - v_1(x_2)$ . The final price of item 1 is then  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = R_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = 0$



with  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ . Thus,  $\forall p \in I_\alpha^4$  with  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$ , we have:

$$\begin{aligned} \|f(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (R_1 - p_1)^2 + (p_2)^2 \\ &= (R_1 + \alpha - v_1(x_1 + x_2))^2 + \alpha^2 \end{aligned} \quad (\text{B.38})$$

Thus,  $\forall p \in I_\alpha^4, g_\Gamma(p) \geq (v_2(x_1 + x_2) - v_2(x_1))^2 > 0$ .

We have shown that:

- $g_\Gamma(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2}) = 0$
- The only possible candidate given by the study of  $I_\alpha^1$  as an element of the frontier prediction of final prices would be  $(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$
- If  $I_\alpha^3 \neq \emptyset$ , the only possible candidate given by the study of  $I_\alpha^3$  as an element of the frontier prediction of final prices would be  $(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$
- If  $I_\alpha^4 \neq \emptyset, \forall p \in I_\alpha^4, g_\Gamma(p) \geq (v_2(x_1 + x_2) - v_2(x_1))^2 > 0$

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})\}$ .

□

**Proof. Sequence convergence:**

Let's first show that:

$$\begin{aligned} p &\in [0, \frac{v_1(x_1 + x_2)}{2}) \times [0, \frac{v_1(x_1 + x_2)}{2}) \\ \implies f_\Gamma(p_1, p_2) &= (\frac{v_1(x_1 + x_2)}{2}, \frac{v_1(x_1 + x_2)}{2}) \end{aligned} \quad (\text{B.39})$$

Suppose  $p \in [0, \frac{v_1(x_1+x_2)}{2}) \times [0, \frac{v_1(x_1+x_2)}{2})$ . As  $p_1 < \frac{v_1(x_1+x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_2)$ ,  $p_2 < \frac{v_1(x_1+x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 + p_2 <$

$v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , player 2 will bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $p_1 < \frac{v_1(x_1+x_2)}{2} \leq v_1(x_1 + x_2) - v_1(x_2)$  and  $p_2 < \frac{v_1(x_1+x_2)}{2} \leq v_1(x_1 + x_2) - v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed. Thus, the price of each item rises symmetrically. Player 1 will stop bidding on both items when their price reaches  $\frac{v_1(x_1+x_2)}{2}$ . Player 2 will then acquire each item for  $\frac{v_1(x_1+x_2)}{2}$ . Thus,  $f_\Gamma(p_1, p_2) = (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$ .

As  $P_0 = (0, 0) < (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$  component-wise and  $P_{t+1} = \frac{1}{t+1}f_\Gamma(P_t) + (1 - \frac{1}{t+1})P_t$ , by a simple induction we have that  $\forall t \in \mathbb{R}_+, P_t \in [0, \frac{v_1(x_1+x_2)}{2}) \times [0, \frac{v_1(x_1+x_2)}{2})$  and, thus,  $\forall t \in \mathbb{R}_+, f_\Gamma(P_t) = (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$ .

The above results directly imply that  $P_t \xrightarrow{t \rightarrow +\infty} (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

### Second sub-sub-sub-case:

$$* \quad v_2(x_1 + x_2) - v_2(x_1) < \frac{v_1(x_1+x_2)}{2}$$

*Proof. Frontier prediction of final prices:*

We decompose  $I_\alpha$  in 3 disjoint sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2), v_2(x_1 + x_2) - v_2(x_1)]\}$
- $I_\alpha^2 = \{(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))\}$
- $I_\alpha^3 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$

such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3$ . We will calculate the infimum of  $g_\Gamma$  on each of these three sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- Suppose  $p \in I_\alpha^1$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_2(x_1 + x_2) - v_2(x_2) > p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1) > p_2 + \eta_1$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , **player 2 will bid on both items if their price is below  $p_2 + \eta_1$** . As  $v_1(x_1 + x_2) - v_1(x_1) > p_2 - \eta_2$ ,  $v_1(x_1 + x_2) - v_1(x_2) > p_1 - \eta_1$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , **player 1 will bid on both items**. Prices will then rise symmetrically. As  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1) > p_2 + \eta_1$ , **player 1 will drop out of the auction when the price of each item reaches  $p_2 + \eta_1$** . Thus, **player 2 will acquire each item for  $p_2 + \eta_1$** . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ . Thus,  $\forall p \in I_\alpha^1$  with  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \tag{B.40}$$

The function  $\alpha \rightarrow \lim_{\eta \rightarrow 0^+} \frac{1}{\eta^2} \int_{\Omega=[0, \eta]^2} (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 d\eta_1 d\eta_2$  admits a minimum for  $\alpha = \frac{v_1(x_1 + x_2)}{2} \geq v_2(x_1 + x_2) - v_2(x_1)$ .

Thus, the only possible candidate given by the study of  $I_\alpha^1$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$

- Suppose  $p \in I_\alpha^2$  and, thus,  $p = (v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_2(x_1 + x_2) - v_2(x_1) < v_2(x_1)$ , **player 2 will always bid on both items if their price is below  $v_2(x_1 + x_2) - v_2(x_1)$** . As  $v_2(x_1 + x_2) - v_2(x_1) < v_1(x_1 + x_2) - v_1(x_2)$ ,  $v_2(x_1 + x_2) - v_2(x_1) < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , **player 1 also always bids on both items if their price is below  $v_2(x_1 + x_2) - v_2(x_1)$** . Thus, the price of each item rises symmetrically. As  $p_2 = v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1)$ , the complementarity surplus of player 2 will be entirely consumed when

the price of item 2 reaches  $p_2$ . Player 1 will then acquire item 2 at price  $p_2$ . Player 1 is then currently exposed of  $p_2 - v_1(x_2)$ . It is then beneficial for player 1 to acquire item 1 if its price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . Player 2 will continue bidding on item 1 if its price is below  $v_2(x_1)$ . Thus, if player 1 stops bidding first on item 1, then its final price is  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding first on item 1, then its final price is  $v_2(x_1)$ . We set  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$  and, thus, the final price of item 1 is  $R_1$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = R_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2$  with  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ . Thus, we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (R_1 - p_1)^2 + (p_2 - p_2)^2 \\ &= (R_1 - p_1)^2 \end{aligned} \quad (\text{B.41})$$

Thus,  $g_\Gamma(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1)) = (R_1 - p_1)^2$

- Suppose  $p \in I_\alpha^3$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > v_2(x_1 + x_2) - v_2(x_1)$  and, thus, player 2 won't bid on item 2. The final price of item 2 is then 0. However, player 1 predicts that item 2 will be sold for  $p_2 - \eta_2 > v_1(x_2)$  and, hence, she thinks that she is exposed of  $p_2 - \eta_2 - v_1(x_2)$ . It is then beneficial for player 1 to acquire item 1 if its price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . Player 2 will however continue to bid on item 1 until its price reaches  $v_2(x_1)$ . Thus, if player 1 stops bidding on item 1 first, then its final price is  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding on item 1 first, then its final price is  $v_2(x_1)$ . We set  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$  and, thus, the final price of item 1 is  $R_1$ . More formally, we have  $\forall p \in I_\alpha^3, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = R_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = 0$  with  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1)]$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (R_1 - p_1)^2 + (p_2)^2 \\ &= (R_1 + \alpha - v_1(x_1 + x_2))^2 + (\alpha)^2 \end{aligned} \quad (\text{B.42})$$

The above function is minimised for  $\alpha = \frac{1}{2} \min(v_1(x_2), v_1(x_1 + x_2) - v_2(x_1)) \leq v_2(x_1 + x_2) - v_2(x_1)$ .

Thus, the only possible candidate given by the study of  $I_\alpha^3$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .

We have shown that:

- $g_\Gamma(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1)) = (R_1 - p_1)^2$
- The only possible candidate given by the study of  $I_\alpha^1$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .
- The only possible candidate given by the study of  $I_\alpha^3$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))\}$ .  $\square$

**Proof. Sequence convergence:**

Let's set  $V_1 = v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1))$ ,  $V_2 = v_2(x_1 + x_2) - v_2(x_1)$  and  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ .

Let's define:

- $S_1 = \{p \in \mathbb{R}_+^2, p_1 \leq V_1, p_2 < V_2\}$
- $S_2 = \{p \in \mathbb{R}_+^2, V_1 < p_1 < R_1, p_1 + p_2 < v_1(x_1 + x_2)\}$
- $S_3 = \{p \in \mathbb{R}_+^2, V_1 < p_1 < R_1, p_2 < V_2, p_1 + p_2 \geq v_1(x_1 + x_2)\}$

Let's show the three following points:

- $\forall p \in S_1, f_{\Gamma}^1(p) = R_1 \geq V_1, f_{\Gamma}^2(p) = V_2$
- $\forall p \in S_2, f_{\Gamma}^1(p) = f_{\Gamma}^2(p) = v_1(x_1 + x_2) - p_1 \leq V_2 < V_1$
- $\forall p \in S_3, f_{\Gamma}^1(p) = f_{\Gamma}^2(p) = 0 < V_2 < V_1$

Suppose  $p \in S_1$ . As  $V_2 < V_1 \leq v_2(x_1)$ , player 2 will always bid on both items if their price is below  $V_2$ . As  $V_2 < v_1(x_1 + x_2) - v_1(x_1)$ ,  $V_2 < V_1 < v_1(x_1 + x_2) - v_1(x_2)$  and  $V_2 < \frac{v_1(x_1 + x_2)}{2}$ , player 1 also always bids on both items if their price is below  $V_2$ . Thus, the price of each item rises symmetrically. As  $V_2 < v_2(x_1)$ , the complementarity surplus of player 2 will be entirely consumed when the price of item 2 reaches  $V_2$ . Player 1 will then acquire item 2 at price  $V_2$ . Player 1 is then currently exposed of  $V_2 - v_1(x_2)$ . It is then beneficial for player 1 to acquire item 1 if its price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . Player 2 will continue bidding on item 1 if its price is below  $v_2(x_1)$ . Thus, if player 1 stops bidding first on item 1, then its final price is  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding first on item 1, then its final price is  $v_2(x_1)$ . The final price of item 1 is then  $R_1$ . Thus, we have  $f_{\Gamma}^1(p) = R_1 \geq V_1, f_{\Gamma}^2(p) = V_2$ .

This proves the first point.

Suppose  $p \in S_2$ . As  $p_2 < V_2, p_1 < v_2(x_1 + x_2) - v_2(x_2)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items until her complementarity surplus is consumed entirely. As  $p_2 < v_1(x_1 + x_2) - v_1(x_1), p_1 < v_1(x_1 + x_2) - v_1(x_2)$  and  $p_1 + p_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items until her complementarity surplus is consumed entirely. As  $p_2 < v_1(x_1 + x_2) - p_1 < V_2 < V_1 < p_1$ , player 1 complementarity surplus will be entirely consumed when the price of item 2 reaches  $v_1(x_1 + x_2) - p_1$ . Player 1 will then drop out of the auction and player 2 will acquire each item for  $v_1(x_1 + x_2) - p_1$ . Thus, we have  $f_{\Gamma}^1(p) = f_{\Gamma}^2(p) = v_1(x_1 + x_2) - p_1 < V_2 < V_1$ .

This proves the second point.

Suppose  $p \in S_3$ . As  $p_1 + p_2 \geq v_1(x_1 + x_2)$ ,  $v_1(x_1) < V_1 < p_1$  and  $v_1(x_2) \leq v_1(x_1 + x_2) - R_1 < p_2$ , player 1 won't bid on any items. Thus, player 2 will acquire both items for zero. Thus,  $f_\Gamma^1(p) = f_\Gamma^2(p) = 0 < V_2 < V_1$ .

This proves the third point.

Regarding sequence  $P_{t+1} = \frac{1}{t+1} f_\Gamma(P_t) + (1 - \frac{1}{t+1}) P_t$ , the three points above directly imply that:

- If  $P_t \in S_1$ , then, as  $f_\Gamma^1(P_t) = R_1$  and  $f_\Gamma^2(P_t) = V_2$ ,  $P_{t+1} \in S_1 \cup S_2 \cup S_3$ .
- If  $P_t \in S_2$ , then, as  $f_\Gamma^1(P_t) = f_\Gamma^2(P_t) = v_1(x_1 + x_2) - p_1 \leq V_2 < V_1$ ,  $P_{t+1} \in S_1 \cup S_2$ .
- If  $P_t \in S_3$ , then, as  $f_\Gamma^1(p) = f_\Gamma^2(p) = 0$ ,  $P_{t+1} \in S_1 \cup S_2 \cup S_3$ .

As  $P_0 = (0, 0) \in S_1$ , we have by induction that  $\forall t \in \mathbb{R}_+$ ,  $P_t \in S_1 \cup S_2 \cup S_3$ .

Moreover,

$$\begin{cases} \text{If } P_t^1 \leq V_1, P_t^1 \in S_1 \text{ and, thus, } f_\Gamma^1(P_t) > V_1 \\ \text{If } P_t^1 > V_1, P_t^1 \in S_2 \cup S_3 \text{ and, thus, } f_\Gamma^1(P_t) < V_1 \end{cases} \quad (\text{B.43})$$

This directly implies that  $P_t^1 \xrightarrow[t \rightarrow +\infty]{} V_1$ .

Let's show that  $P_t^2 \xrightarrow[t \rightarrow +\infty]{} V_2$ , i.e.,  $\forall \eta > 0$ ,  $\exists t_0 \in \mathbb{R}_+$ ,  $\forall t \geq t_0$ ,  $P_t^2 \in [V_2 - \eta, V_2]$ .

Let's fix  $\eta > 0$ .

Let's fix  $t_1 = \lceil 2(\frac{V_2}{\eta} - 1) \rceil$ .

As  $P_t^1 \xrightarrow[t \rightarrow +\infty]{} V_1$ ,  $\exists t_2 \in \mathbb{R}_+$ ,  $\forall t \geq t_2$ ,  $P_t^1 \in [V_1 - \frac{\eta}{2}, V_1 + \frac{\eta}{2}]$ . Let's fix such a  $t_2$ .

Suppose  $t \geq \max(t_1, t_2)$ . If  $P_t^2 < V_2 - \frac{\eta}{2}$ , then  $P_t \in S_1 \cup S_2$  and, thus,  $f_\Gamma^2(P_t) \geq V_2 - \frac{\eta}{2}$ . Hence,  $\exists t_3 > \max(t_1, t_2)$ ,  $P_t^2 \in [V_2 - \eta, V_2)$ .

Let's fix such a  $t_3$ .

We will show by induction that  $\forall t \geq t_3$ ,  $P_t^2 \in [V_2 - \eta, V_2)$  and, hence,  $P_t^2 \xrightarrow{t \rightarrow +\infty} V_2$ .

Suppose  $t > t_3$  and  $P_t^2 \in [V_2 - \eta, V_2)$ . Three cases can occur.

- If  $P_t \in S_1$ , then  $f_\Gamma^2(P_t) = V_2$  and, thus,  $P_{t+1}^2 \in [V_2 - \eta, V_2)$ .
- If  $P_t \in S_2$ , then  $f_\Gamma^2(P_t) \geq V_2 - \frac{\eta}{2}$  and, thus,  $P_{t+1}^2 \in [V_2 - \eta, V_2)$ .
- If  $P_t \in S_3$ , then  $P_t^2 \geq V_2 - \frac{\eta}{2}$  and  $f_\Gamma^2(P_t) = 0$ . We have:

$$\begin{aligned}
 P_{t+1}^2 &= \frac{t}{t+1} P_t^2 \geq \frac{t}{t+1} (V_2 - \frac{\eta}{2}) \\
 &\geq \frac{2(\frac{V_2}{\eta} - 1)}{2(\frac{V_2}{\eta} - 1) + 1} (V_2 - \frac{\eta}{2}) \quad (\text{as } t > t_3 > t_1 = \lceil 2(\frac{V_2}{\eta} - 1) \rceil) \\
 &\geq \frac{(\frac{V_2}{\eta} - 1)}{\frac{1}{\eta}(2V_2 - \eta)} (2V_2 - \eta) \\
 &\geq V_2 - \eta
 \end{aligned}
 \tag{B.44}$$

Thus,  $P_{t+1}^2 \in [V_2 - \eta, V_2)$ .

We have shown by induction that  $\forall t \geq t_3$ ,  $P_t^2 \in [V_2 - \eta, V_2)$ . Hence,  $P_t^2 \xrightarrow{t \rightarrow +\infty} V_2$ .

We have shown that  $P_t \xrightarrow{t \rightarrow +\infty} (V_1, V_2)$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

## B. Second sub-sub-case:



$$* v_1(x_1) > \frac{v_1(x_1+x_2)}{2}$$

**First sub-sub-sub-case:**

$$* v_2(x_1 + x_2) - v_2(x_1) \geq v_1(x_1 + x_2) - v_1(x_1)$$

**Proof. Frontier prediction of final prices:**

We decompose  $I_\alpha$  in 2 disjoint sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2), v_1(x_1 + x_2) - v_1(x_1)]\}$
- $I_\alpha^2 = \{(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$

such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2$ . We will calculate the infimum of  $g_\Gamma$  on each of these two sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- Suppose  $p \in I_\alpha^1$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_1(x_1 + x_2) - v_1(x_1)$ . As  $v_2(x_1 + x_2) - v_2(x_2) > p_1 - \eta_1 > v_1(x_1 + x_2) - v_1(x_1)$ ,  $v_2(x_1 + x_2) - v_2(x_1) \geq v_1(x_1 + x_2) - v_1(x_1) > p_2 + \eta_1$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items if their price is below  $p_2 + \eta_1$ . As  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items until her complementarity surplus is consumed. As  $p_1 - \eta_1 > v_1(x_1 + x_2) - v_1(x_1) > p_2 + \eta_1$ , player 1 will drop out of the auction when the price of each item reaches  $p_2 + \eta_1$ . Thus, player 2 will acquire each item for  $p_2 + \eta_1$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ . Thus, we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \tag{B.45}$$

Thus,  $\forall p \in I_\alpha^1, g_\Gamma(p) > 0$ .

- Suppose  $p \in I_\alpha^2$  and, thus,  $p = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ . As  $p_1 - \eta_1 < v_2(x_1)$  and  $p_2 - \eta_2 < v_2(x_1 + x_2) - v_2(x_1)$ , player 2 will bid on both items until her complementarity surplus is consumed. As  $p_1 - \eta_1 < v_1(x_1)$  and  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus is consumed. Thus, the price of each item rises symmetrically. As  $v_1(x_1 + x_2) - v_1(x_1) \leq v_1(x_1)$ , player 1 complementarity surplus is entirely consumed by item 2. Thus, player 2 acquires item 2 for  $v_1(x_1 + x_2) - v_1(x_1)$ . Both players continue bidding on item 1. As  $v_1(x_1) \leq v_2(x_1)$ , player 2 acquires item 1 for  $v_1(x_1)$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = p_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2$ . Thus, we have:

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = 0 \quad (\text{B.46})$$

Thus,  $g_\Gamma(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) = 0$ .

We have shown that:

- $\forall p \in I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2), v_1(x_1 + x_2) - v_1(x_1)]\}, g_\Gamma(p) > 0$
- $g_\Gamma(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) = 0$

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$ .  $\square$

*Proof. Sequence convergence:*

Let's first show that:

$$\begin{aligned} p &\in [0, v_1(x_1)] \times [0, v_1(x_1 + x_2) - v_1(x_1)] \\ \implies f_\Gamma(p_1, p_2) &= (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) \end{aligned} \quad (\text{B.47})$$

**Suppose**  $p \in [0, v_1(x_1)] \times [0, v_1(x_1 + x_2) - v_1(x_1)]$ . **As**  $v_1(x_1 + x_2) - v_1(x_1) < v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$ ,  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_1)$  and  $v_1(x_1 + x_2) - v_1(x_1) \leq \frac{v_2(x_1 + x_2)}{2}$ , **player 2 will bid on both items if their price is below**  $v_1(x_1 + x_2) - v_1(x_1)$ . **As**  $p_1 < v_1(x_1)$  and  $p_2 < v_1(x_1 + x_2) - v_1(x_1) < v_1(x_1)$ , **player 1 will bid on both items until her complementarity surplus is entirely consumed by item 2.** Thus, the price of each item rises symmetrically. **Player 1 will stop bidding on item 2 when its price reaches**  $v_1(x_1 + x_2) - v_1(x_1)$  and will continue to bid on item 1 until its price reaches  $v_1(x_1)$ . **Player 2 will then acquire item 2 for**  $v_1(x_1 + x_2) - v_1(x_1)$  and will continue bidding on item 1 until its price reaches  $v_2(x_1 + x_2) - v_2(x_2)$ . **As**  $v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$ , **player 2 acquires item 1 for**  $v_1(x_1)$ . Thus,  $f_\Gamma(p_1, p_2) = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ .

**As**  $P_0 = (0, 0) < (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$  component-wise and  $P_{t+1} = \frac{1}{t+1} f_\Gamma(P_t) + (1 - \frac{1}{t+1}) P_t$ , **by a simple induction we have that**  $\forall t \in \mathbb{R}_+, P_t \in [0, v_1(x_1)] \times [0, v_1(x_1 + x_2) - v_1(x_1)]$  and, thus,  $\forall t \in \mathbb{R}_+, f_\Gamma(P_t) = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ .

The above results directly imply that  $P_t \xrightarrow[t \rightarrow +\infty]{} (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{\min*}$ .

□

**Second sub-sub-sub-case:**

$$* \quad v_2(x_1 + x_2) - v_2(x_1) < v_1(x_1 + x_2) - v_1(x_1)$$

*Proof.* **Frontier prediction of final prices:**

We decompose  $I_\alpha$  in 3 disjoint sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2), v_2(x_1 + x_2) - v_2(x_1)]\}$
- $I_\alpha^2 = \{(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))\}$
- $I_\alpha^3 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$

such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3$ . We will calculate the infimum of  $g_\Gamma$  on each of these three sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- Suppose  $p \in I_\alpha^1$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_2(x_1 + x_2) - v_2(x_2) > p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1) > p_2 + \eta_1$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items if their price is below  $p_2 + \eta_1$ . As  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items until her complementarity surplus is consumed. As  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1) > p_2 + \eta_1$ , player 1 will drop out of the auction when the price of each item reaches  $p_2 + \eta_1$ . Thus, player 2 will acquire each item for  $p_2 + \eta_1$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ . Thus, we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \tag{B.48}$$

The function  $\alpha \rightarrow \lim_{\eta \rightarrow 0^+} \frac{1}{\eta^2} \int_{\Omega=[0, \eta]^2} (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 d\eta_1 d\eta_2$  admits a minimum for  $\alpha = \frac{v_1(x_1 + x_2)}{2} \geq v_2(x_1 + x_2) - v_2(x_1)$ .

Thus, the only possible candidate given by the study of  $I_\alpha^1$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .

- Suppose  $p \in I_\alpha^2$  and, thus,  $p = (v_1(x_1+x_2)-(v_2(x_1+x_2)-v_2(x_1)), v_2(x_1+x_2) - v_2(x_1))$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_2(x_1 + x_2) - v_2(x_1) < v_2(x_1)$ , player 2 will always bid on both items if their price is below  $v_2(x_1+x_2)-v_2(x_1)$ . As  $v_2(x_1+x_2)-v_2(x_1) < v_1(x_1+x_2)-v_1(x_2)$ ,  $v_2(x_1 + x_2) - v_2(x_1) < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 also always bids on both items if their price is below  $v_2(x_1+x_2)-v_2(x_1)$ . Thus, the price of each item rises symmetrically. As  $p_2 = v_2(x_1+x_2)-v_2(x_1)$  and  $p_1 - \eta_1 > v_2(x_1+x_2)-v_2(x_1)$ , the complementarity surplus of player 2 will be entirely consumed when the price of item 2 reaches  $p_2$ . Player 1 will then acquire item 2 at price  $p_2$ . Player 1 is then currently exposed of  $p_2 - v_1(x_2)$ . It is then beneficial for player 1 to acquire item 1 if its price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . Player 2 will continue bidding on item 1 if its price is below  $v_2(x_1)$ . Thus, if player 1 stops bidding first on item 1, then its final price is  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding first on item 1, then its final price is  $v_2(x_1)$ . We set  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$  and, thus, the final price of item 1 is  $R_1$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = R_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2$  with  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ . Thus, we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (R_1 - p_1)^2 + (p_2 - p_2)^2 \\ &= (R_1 - p_1)^2 \end{aligned} \quad (\text{B.49})$$

Thus,

$$g_\Gamma(v_1(x_1+x_2)-(v_2(x_1+x_2)-v_2(x_1)), v_2(x_1+x_2)-v_2(x_1)) = (R_1 - p_1)^2$$

- Suppose  $p \in I_\alpha^3$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > v_2(x_1 + x_2) - v_2(x_1)$  and, thus, player 2 won't bid on item 2. The final price of item 2 is then 0. However, player 1 predicts that item 2 will be sold for  $p_2 - \eta_2 > v_1(x_2)$  and, hence, she thinks that she is exposed of  $p_2 - \eta_2 - v_1(x_2)$ . It

is then beneficial for player 1 to acquire item 1 if its price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . Player 2 will however continue to bid on item 1 until its price reaches  $v_2(x_1)$ . Thus, if player 1 stops bidding on item 1 first, then its final price is  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding on item 1 first, then its final price is  $v_2(x_1)$ . We set  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$  and, thus, the final price of item 1 is  $R_1$ . More formally, we have  $\forall p \in I_\alpha^3, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = R_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = 0$  with  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in (v_2(x_1 + x_2) - v_2(x_1), v_1(x_1 + x_2) - v_1(x_1)]$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (R_1 - p_1)^2 + (p_2)^2 \\ &= (R_1 + \alpha - v_1(x_1 + x_2))^2 + (\alpha)^2 \end{aligned} \quad (\text{B.50})$$

The above function is minimised for  $\alpha = \frac{1}{2} \min(v_1(x_2), v_1(x_1 + x_2) - v_2(x_1)) \leq v_2(x_1 + x_2) - v_2(x_1)$ .

Thus, the only possible candidate given by the study of  $I_\alpha^3$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .

We have shown that:

- $g_\Gamma(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1)) = (R_1 - p_1)^2$
- The only possible candidate given by the study of  $I_\alpha^1$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .
- The only possible candidate given by the study of  $I_\alpha^3$  as an element of the frontier prediction of final prices would be  $(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))$ .

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1)), v_2(x_1 + x_2) - v_2(x_1))\}$ .

□

**Proof. Sequence convergence:**

Let's set  $V_1 = v_1(x_1 + x_2) - (v_2(x_1 + x_2) - v_2(x_1))$ ,  $V_2 = v_2(x_1 + x_2) - v_2(x_1)$  and  $R_1 = \min(v_2(x_1), v_1(x_1 + x_2) - v_1(x_2))$ .

Let's define:

- $S_1 = \{p \in \mathbb{R}_+^2, p_1 \leq V_1, p_2 < V_2\}$
- $S_2 = \{p \in \mathbb{R}_+^2, V_1 < p_1 < R_1, p_1 + p_2 < v_1(x_1 + x_2)\}$
- $S_3 = \{p \in \mathbb{R}_+^2, V_1 < p_1 < R_1, p_2 < V_2, p_1 + p_2 \geq v_1(x_1 + x_2)\}$

Let's show the three following points:

- $\forall p \in S_1, f_{\Gamma}^1(p) = R_1 \geq V_1, f_{\Gamma}^2(p) = V_2$
- $\forall p \in S_2, f_{\Gamma}^1(p) = f_{\Gamma}^2(p) = v_1(x_1 + x_2) - p_1 \leq V_2 < V_1$
- $\forall p \in S_3, f_{\Gamma}^1(p) = f_{\Gamma}^2(p) = 0 < V_2 < V_1$

Suppose  $p \in S_1$ . As  $V_2 < V_1 \leq v_2(x_1)$ , player 2 will always bid on both items if their price is below  $V_2$ . As  $V_2 < v_1(x_1 + x_2) - v_1(x_1)$ ,  $V_2 < V_1 < v_1(x_1 + x_2) - v_1(x_2)$  and  $V_2 < \frac{v_1(x_1 + x_2)}{2}$ , player 1 also always bids on both items if their price is below  $V_2$ . Thus, the price of each item rises symmetrically. As  $V_2 < v_2(x_1)$ , the complementarity surplus of player 2 will be entirely consumed when the price of item 2 reaches  $V_2$ . Player 1 will then acquire item 2 at price  $V_2$ . Player 1 is then currently exposed of  $V_2 - v_1(x_2)$ . It is then beneficial for player 1 to acquire item 1 if its price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . Player 2 will continue bidding on item 1 if its price is below  $v_2(x_1)$ . Thus, if player 1 stops bidding first on item 1, then its final price is  $v_1(x_1 + x_2) - v_1(x_2)$ . If player 2 stops bidding first on item 1, then its final price is  $v_2(x_1)$ . The final price of item 1 is then  $R_1$ . Thus, we have  $f_{\Gamma}^1(p) = R_1 \geq V_1, f_{\Gamma}^2(p) = V_2$ .

This proves the first point.

Suppose  $p \in S_2$ . As  $p_2 < V_2$ ,  $p_1 < v_2(x_1 + x_2) - v_2(x_2)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items until her complementarity surplus is consumed entirely. As  $p_2 < v_1(x_1 + x_2) - v_1(x_1)$ ,  $p_1 < v_1(x_1 + x_2) - v_1(x_2)$  and  $p_1 + p_2 < v_1(x_1 + x_2)$ , player 1 will bid on both items until her complementarity surplus is consumed entirely. As  $p_2 < v_1(x_1 + x_2) - p_1 < V_2 < V_1 < p_1$ , player 1 complementarity surplus will be entirely consumed when the price of item 2 reaches  $v_1(x_1 + x_2) - p_1$ . Player 1 will then drop out of the auction and player 2 will acquire each item for  $v_1(x_1 + x_2) - p_1$ . Thus, we have  $f_\Gamma^1(p) = f_\Gamma^2(p) = v_1(x_1 + x_2) - p_1 < V_2 < V_1$ .

This proves the second point.

Suppose  $p \in S_3$ . As  $p_1 + p_2 \geq v_1(x_1 + x_2)$ ,  $v_1(x_1) < V_1 < p_1$  and  $v_1(x_2) \leq v_1(x_1 + x_2) - R_1 < p_2$ , player 1 won't bid on any items. Thus, player 2 will acquire both items for zero. Thus,  $f_\Gamma^1(p) = f_\Gamma^2(p) = 0 < V_2 < V_1$ .

This proves the third point.

Regarding sequence  $P_{t+1} = \frac{1}{t+1} f_\Gamma(P_t) + (1 - \frac{1}{t+1}) P_t$ , the three points above directly imply that:

- If  $P_t \in S_1$ , then, as  $f_\Gamma^1(P_t) = R_1$  and  $f_\Gamma^2(P_t) = V_2$ ,  $P_{t+1} \in S_1 \cup S_2 \cup S_3$ .
- If  $P_t \in S_2$ , then, as  $f_\Gamma^1(P_t) = f_\Gamma^2(P_t) = v_1(x_1 + x_2) - p_1 \leq V_2 < V_1$ ,  $P_{t+1} \in S_1 \cup S_2$ .
- If  $P_t \in S_3$ , then, as  $f_\Gamma^1(p) = f_\Gamma^2(p) = 0$ ,  $P_{t+1} \in S_1 \cup S_2 \cup S_3$ .

As  $P_0 = (0, 0) \in S_1$ , we have by induction that  $\forall t \in \mathbb{R}_+$ ,  $P_t \in S_1 \cup S_2 \cup S_3$ .

Moreover,

$$\begin{cases} \text{If } P_t^1 \leq V_1, P_t^1 \in S_1 \text{ and, thus, } f_\Gamma^1(P_t) > V_1 \\ \text{If } P_t^1 > V_1, P_t^1 \in S_2 \cup S_3 \text{ and, thus, } f_\Gamma^1(P_t) < V_1 \end{cases} \quad (\text{B.51})$$



This directly implies that  $P_t^1 \xrightarrow{t \rightarrow +\infty} V_1$ .

Let's show that  $P_t^2 \xrightarrow{t \rightarrow +\infty} V_2$ , i.e.,  $\forall \eta > 0, \exists t_0 \in \mathbb{R}_+, \forall t \geq t_0, P_t^2 \in [V_2 - \eta, V_2)$ .

Let's fix  $\eta > 0$ .

Let's fix  $t_1 = \lceil 2(\frac{V_2}{\eta} - 1) \rceil$ .

As  $P_t^1 \xrightarrow{t \rightarrow +\infty} V_1, \exists t_2 \in \mathbb{R}_+, \forall t \geq t_2, P_t^1 \in [V_1 - \frac{\eta}{2}, V_1 + \frac{\eta}{2}]$ . Let's fix such a  $t_2$ .

Suppose  $t \geq \max(t_1, t_2)$ . If  $P_t^2 < V_2 - \frac{\eta}{2}$ , then  $P_t \in S_1 \cup S_2$  and, thus,  $f_\Gamma^2(P_t) \geq V_2 - \frac{\eta}{2}$ . Hence,  $\exists t_3 > \max(t_1, t_2), P_t^2 \in [V_2 - \eta, V_2)$ .

Let's fix such a  $t_3$ .

We will show by induction that  $\forall t \geq t_3, P_t^2 \in [V_2 - \eta, V_2)$  and, hence,  $P_t^2 \xrightarrow{t \rightarrow +\infty} V_2$ .

Suppose  $t > t_3$  and  $P_t^2 \in [V_2 - \eta, V_2)$ . Three cases can occur.

- If  $P_t \in S_1$ , then  $f_\Gamma^2(P_t) = V_2$  and, thus,  $P_{t+1}^2 \in [V_2 - \eta, V_2)$ .
- If  $P_t \in S_2$ , then  $f_\Gamma^2(P_t) \geq V_2 - \frac{\eta}{2}$  and, thus,  $P_{t+1}^2 \in [V_2 - \eta, V_2)$ .
- If  $P_t \in S_3$ , then  $P_t^2 \geq V_2 - \frac{\eta}{2}$  and  $f_\Gamma^2(P_t) = 0$ . We have:

$$\begin{aligned}
 P_{t+1}^2 &= \frac{t}{t+1} P_t^2 \geq \frac{t}{t+1} (V_2 - \frac{\eta}{2}) \\
 &\geq \frac{2(\frac{V_2}{\eta} - 1)}{2(\frac{V_2}{\eta} - 1) + 1} (V_2 - \frac{\eta}{2}) \quad (\text{as } t > t_1 = \lceil 2(\frac{V_2}{\eta} - 1) \rceil) \\
 &\geq \frac{(\frac{V_2}{\eta} - 1)}{\frac{1}{\eta}(2V_2 - \eta)} (2V_2 - \eta) \\
 &\geq V_2 - \eta
 \end{aligned}$$

(B.52)

Thus,  $P_{t+1}^2 \in [V_2 - \eta, V_2)$ .

We have shown by induction that  $\forall t \geq t_3, P_t^2 \in [V_2 - \eta, V_2)$ . Hence,  

$$P_t^2 \xrightarrow{t \rightarrow +\infty} V_2.$$

We have shown that  $P_t \xrightarrow{t \rightarrow +\infty} (V_1, V_2)$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

### 3. Third disposition:

$$\begin{cases} v_1(x_1) \geq v_2(x_1) \\ v_1(x_2) \geq v_2(x_2) \\ v_1(x_1 + x_2) \leq v_2(x_1 + x_2) \end{cases} \quad (\text{B.53})$$

*Proof.* **Set of smallest overestimated final prices:**

The following result is shared amongst all the cases of the third disposition. In this proof, we are going to calculate the set of smallest overestimated final prices  $O_\Gamma^{min}$ .

Let's set  $O_\Gamma^1 = \{p, v_1(x_1 + x_2) = p_1 + p_2, v_1(x_1) \leq p_1, v_1(x_2) \leq p_2\}$ . Let's show that  $O_\Gamma^{min} = O_\Gamma^1$  in three steps:

- $O_\Gamma^1 \subset O_\Gamma$
- $\forall p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 < v_1(x_1 + x_2)$ , then  $p \notin O_\Gamma$ . Thus,  $O_\Gamma^1 \subset O_\Gamma^{min}$ .
- $\forall p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 = v_1(x_1 + x_2)$ , if  $p_1 < v_1(x_1)$  or  $p_2 < v_1(x_2)$  then  $p \notin O_\Gamma$ . Thus,  $O_\Gamma^1 = O_\Gamma^{min}$ .

If  $p \in O_\Gamma^1$ , then  $f_\Gamma^1(p) = f_\Gamma^2(p) = 0$  as player 1 predicts that she will get a negative payoff if she acquires any item. Thus,  $p \in O_\Gamma$ . This verifies the first step.

Suppose  $p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 < v_1(x_1 + x_2)$ . Then, we have:

- If  $p_1 \in [0, v_2(x_1))$ , then, regardless of price  $p_2$ ,  $f_\Gamma^1(p) \geq v_2(x_1)$  as both players has their stand-alone value for item 1 greater or equal to  $v_2(x_1)$ . Thus,  $p \notin O_\Gamma$ .
- If  $p_1 \in [v_2(x_1), v_1(x_1))$ , we have two different cases to study.
  - If  $p_2 \geq v_1(x_1 + x_2) - v_1(x_1)$ , then player 1 will bid only on item 1 until its price is  $v_1(x_1)$ . As  $p_2 < v_2(x_1 + x_2) - v_2(x_1)$ ,  $p_1 < v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items. Given that player 2 has acquired item 2 and doesn't want to be exposed, she will bid on item 1 if its price is below  $v_2(x_1 + x_2) - v_2(x_2)$ . Thus,  $f_\Gamma^1(p) = v_1(x_1) > p_1$  and  $p \notin O_\Gamma$ .
  - Otherwise  $p_2 < v_1(x_1 + x_2) - v_1(x_1)$ , player 1 will then bid on both items until consuming entirely her complementarity surplus. As  $p_2 < v_2(x_1 + x_2) - v_2(x_1)$ ,  $p_1 < v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , player 2 will also bid on both items. As  $v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , the auction will end with  $f_\Gamma^1(p) + f_\Gamma^2(p) = v_1(x_1 + x_2) > p_1 + p_2$ . Thus,  $p \notin O_\Gamma$ .
- If  $p_1 \in [v_1(x_1), v_1(x_1 + x_2) - v_1(x_2))$ , player 1 will then bid on both items until consuming entirely her complementarity surplus. As  $p_2 < v_2(x_1 + x_2) - v_2(x_1)$ ,  $p_1 < v_1(x_1 + x_2) - v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , player 2 will also bid on both items. The auction will end with  $f_\Gamma^1(p) + f_\Gamma^2(p) = v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ . Thus,  $p \notin O_\Gamma$ .
- If  $p_1 \in [v_1(x_1 + x_2) - v_1(x_2), v_1(x_1 + x_2))$ , then player 1 will only bid on item 2. If  $v_2(x_1 + x_2) - v_2(x_2) < v_1(x_1 + x_2)$  and  $p_1 \geq v_2(x_1 + x_2) - v_2(x_2)$ , then  $p_2 < v_2(x_2)$  and  $f_\Gamma^2(p) = v_2(x_2)$ . Thus,  $p \notin O_\Gamma$ . Otherwise,  $p_1 < \min(v_1(x_1 + x_2), v_2(x_1 + x_2) - v_2(x_2))$  and player 2 will bid on item 2 until it has consumed entirely her complementarity surplus. Thus,  $f_\Gamma^2(p) > p_2$  and  $p \notin O_\Gamma$ .

This verifies the second step.

Suppose  $p \in \mathbb{R}_+^2$  such that  $p_1 + p_2 = v_1(x_1 + x_2)$  and  $p_1 < v_1(x_1)$ . As  $p_2 > v_1(x_1 + x_2) - v_1(x_1)$ , player 1 will only bid on item 1. If  $v_2(x_1 + x_2) - v_2(x_1) < v_1(x_1 + x_2)$  and  $p_2 \geq v_2(x_1 + x_2) - v_2(x_1)$ , then  $p_1 < v_2(x_1)$  and  $f_\Gamma^1(p) = v_2(x_1) > p_1$ . Thus,  $p \notin O_\Gamma$ . Otherwise,  $p_2 < \min(v_1(x_1 + x_2), v_2(x_1 + x_2) - v_2(x_1))$ . As  $p_1 < v_2(x_1 + x_2) - v_2(x_2)$ ,  $p_2 < v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 + p_2 < v_2(x_1 + x_2)$ , player 2 will bid on both items.

She will acquire item 2 for a price of zero and, then, bid on item 1 until she has consumed entirely her complementarity surplus. Thus,  $f_{\Gamma}^1(p) > p_1$  and  $p \notin O_{\Gamma}$ .

Symmetrically, we obtain the same result if  $p_2 < v_1(x_2)$ .

This verifies the third step.

Thus, we have shown that:

$$O_{\Gamma}^{min} = \{p, v_1(x_1 + x_2) = p_1 + p_2, v_1(x_1) \leq p_1, v_1(x_2) \leq p_2\} \quad (\text{B.54})$$

□

a) **First case:**

$$* v_1(x_1) \leq \frac{v_1(x_1+x_2)}{2} \text{ and } v_1(x_2) \leq \frac{v_1(x_1+x_2)}{2}$$

**Proof. Frontier prediction of final prices:**

In this proof, we are going to calculate the frontier prediction of final prices and, thus, calculate the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_{\Gamma}|_{O_{\Gamma}^{min}}$  for any  $\gamma > 0$ .

We can rewrite  $O_{\Gamma}^{min}$  as  $I_{\alpha} = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2); v_1(x_1 + x_2) - v_1(x_1)]\}$ . From there, we are going to calculate the set of smallest overestimated final prices which are  $\gamma$  close to the infimum of  $g_{\Gamma}|_{O_{\Gamma}^{min}}$  for any  $\gamma > 0$ . To do so, we need to obtain the closed-form of  $f_{\Gamma}$  for every slight underestimation of elements in  $I_{\alpha}$ . We decompose  $I_{\alpha}$  in 3 disjoint sets

- $I_{\alpha}^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2); \frac{v_1(x_1+x_2)}{2}]\}$
- $I_{\alpha}^2 = \{(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})\}$
- $I_{\alpha}^3 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (\frac{v_1(x_1+x_2)}{2}, v_1(x_1 + x_2) - v_1(x_1)]\}$

such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2 \cup I_\alpha^3$ . We will calculate the infimum of  $g_\Gamma$  on each of these three sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- Suppose  $p \in I_\alpha^1$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > \frac{v_1(x_1+x_2)}{2}$ . As  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , player 2 will always bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 also always bids on both items. Thus, the price of each item rises symmetrically. As  $p_2 + \eta_1 < \frac{v_1(x_1+x_2)}{2}$  and  $p_1 - \eta_1 > \frac{v_1(x_1+x_2)}{2}$ , the complementarity surplus of player 1 will be entirely consumed when the price of item 2 reaches  $p_2 + \eta_1$ . Player 1 will then drop out of the auction and the final price of each item will be  $p_2 + \eta_1$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in [v_1(x_2), \frac{v_1(x_1+x_2)}{2})$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \quad (\text{B.55})$$

Thus,  $\forall p \in I_\alpha^1, g_\Gamma(p) > 0$ .

- Suppose  $p \in I_\alpha^2$  and, thus,  $p = (\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ . As  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , player 2 will always bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $v_1(x_1 + x_2) - v_1(x_1) \geq \frac{v_1(x_1+x_2)}{2}$  and  $v_1(x_1 + x_2) - v_1(x_2) \geq \frac{v_1(x_1+x_2)}{2}$ , player 1 will always bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $p_1 - \eta_1 < \frac{v_1(x_1+x_2)}{2}$  and  $p_2 - \eta_2 < \frac{v_1(x_1+x_2)}{2}$ , both players will always bid on both items and, thus, the price of each item will rise symmetrically until it reaches  $\frac{v_1(x_1+x_2)}{2}$ . Then, player 1 will drop out of the auction. Player 2 acquires each item at a

price  $p_1 = p_2 = \frac{v_1(x_1+x_2)}{2}$ . More formally,  $\exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_1 = p_2$ . Thus, we have

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = 0 \quad (\text{B.56})$$

Thus,  $g_\Gamma(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2}) = 0$ .

- Suppose  $p \in I_\alpha^3$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_2$  is chosen such that  $p_2 - \eta_2 > \frac{v_1(x_1+x_2)}{2}$ . As  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $\frac{v_1(x_1+x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , player 2 will always bid on both items if their price is below  $\frac{v_1(x_1+x_2)}{2}$ . As  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 also always bids on both items. Thus, the price of each item rises symmetrically. As  $p_1 + \eta_2 < \frac{v_1(x_1+x_2)}{2}$  and  $p_2 - \eta_2 > \frac{v_1(x_1+x_2)}{2}$ , the complementarity surplus of player 1 will be entirely consumed when the price of item 1 reaches  $p_1 + \eta_2$ . Player 1 will then drop out of the auction and the final price of each item will be  $p_1 + \eta_2$ . More formally,  $\forall p \in I_\alpha^3, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_1 + \eta_2$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in (\frac{v_1(x_1+x_2)}{2}, v_1(x_1 + x_2) - v_1(x_1)]$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_1 + \eta_2 - p_1)^2 + (p_1 + \eta_2 - p_2)^2 \\ &= (v_1(x_1 + x_2) - 2\alpha + \eta_2)^2 + \eta_2^2 \end{aligned} \quad (\text{B.57})$$

Thus,  $\forall p \in I_\alpha^3, g_\Gamma(p) > 0$ .

We have shown that:

- $\forall p \in I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2), \frac{v_1(x_1+x_2)}{2}]\}$ ,  $g_\Gamma(p) > 0$ .
- $g_\Gamma(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2}) = 0$
- $\forall p \in I_\alpha^3 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in (\frac{v_1(x_1+x_2)}{2}, v_1(x_1 + x_2) - v_1(x_1))\}$ ,  $g_\Gamma(p) > 0$ .

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(\frac{v_1(x_1+x_2)}{2}, \frac{v_1(x_1+x_2)}{2})\}$ .  $\square$

*Proof. Sequence convergence:*

Let's first show that:

$$\begin{aligned} p &\in \left[0, \frac{v_1(x_1 + x_2)}{2}\right) \times \left[0, \frac{v_1(x_1 + x_2)}{2}\right) \\ \implies f_\Gamma(p_1, p_2) &= \left(\frac{v_1(x_1 + x_2)}{2}, \frac{v_1(x_1 + x_2)}{2}\right) \end{aligned} \quad (\text{B.58})$$

Suppose  $p \in \left[0, \frac{v_1(x_1 + x_2)}{2}\right) \times \left[0, \frac{v_1(x_1 + x_2)}{2}\right)$ . As  $p_1 < \frac{v_1(x_1 + x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_2)$ ,  $p_2 < \frac{v_1(x_1 + x_2)}{2} \leq v_2(x_1 + x_2) - v_2(x_1)$  and  $p_1 + p_2 < v_1(x_1 + x_2) \leq v_2(x_1 + x_2)$ , player 2 will bid on both items if their price is below  $\frac{v_1(x_1 + x_2)}{2}$ . As  $p_1 < \frac{v_1(x_1 + x_2)}{2} \leq v_1(x_1 + x_2) - v_1(x_2)$  and  $p_2 < \frac{v_1(x_1 + x_2)}{2} \leq v_1(x_1 + x_2) - v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed. Thus, the price of each item rises symmetrically. Player 1 will stop bidding on both items when their price reaches  $\frac{v_1(x_1 + x_2)}{2}$ . Player 2 will then acquire each item for  $\frac{v_1(x_1 + x_2)}{2}$ . Thus,  $f_\Gamma(p_1, p_2) = \left(\frac{v_1(x_1 + x_2)}{2}, \frac{v_1(x_1 + x_2)}{2}\right)$ .

As  $P_0 = (0, 0) < \left(\frac{v_1(x_1 + x_2)}{2}, \frac{v_1(x_1 + x_2)}{2}\right)$  component-wise and  $P_{t+1} = \frac{1}{t+1}f_\Gamma(P_t) + \left(1 - \frac{1}{t+1}\right)P_t$ , by a simple induction we have that  $\forall t \in \mathbb{R}_+$ ,  $P_t \in \left[0, \frac{v_1(x_1 + x_2)}{2}\right) \times \left[0, \frac{v_1(x_1 + x_2)}{2}\right)$  and, thus,  $\forall t \in \mathbb{R}_+$ ,  $f_\Gamma(P_t) = \left(\frac{v_1(x_1 + x_2)}{2}, \frac{v_1(x_1 + x_2)}{2}\right)$ .

The above results directly imply that  $P_t \xrightarrow{t \rightarrow +\infty} \left(\frac{v_1(x_1 + x_2)}{2}, \frac{v_1(x_1 + x_2)}{2}\right)$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

b) **Second case:**

$$* \quad v_1(x_1) > \frac{v_1(x_1 + x_2)}{2}$$

*Proof. Frontier prediction of final prices:*

We decompose  $I_\alpha$  in 2 disjoints sets

- $I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2); v_1(x_1 + x_2) - v_1(x_1)]\}$
- $I_\alpha^2 = \{(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$

such that  $I_\alpha = I_\alpha^1 \cup I_\alpha^2$ . We will calculate the infimum of  $g_\Gamma$  on each of these two sets and, then, determine the frontier prediction of final prices  $O_\Gamma^{min*}$ . We have:

- Suppose  $p \in I_\alpha^1$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ .  $\eta_1$  is chosen such that  $p_1 - \eta_1 > v_1(x_1)$ . As  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $v_1(x_1 + x_2) - v_1(x_1) \leq \frac{v_2(x_1 + x_2)}{2}$ , player 2 will always bid on both items if their price is below  $v_1(x_1 + x_2) - v_1(x_1)$ . As  $p_1 - \eta_1 < v_1(x_1 + x_2) - v_1(x_2)$ ,  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 + p_2 - \eta_1 - \eta_2 < v_1(x_1 + x_2)$ , player 1 also always bids on both items. Thus, the price of each item rises symmetrically. As  $p_2 + \eta_1 < v_1(x_1 + x_2) - v_1(x_1)$  and  $p_1 - \eta_1 > v_1(x_1 + x_2) - v_1(x_1)$ , the complementarity surplus of player 1 will be entirely consumed when the price of item 2 reaches  $p_2 + \eta_1$ . Player 1 will then drop out of the auction and the final price of each item will be  $p_2 + \eta_1$ . More formally,  $\forall p \in I_\alpha^1, \exists \eta > 0$  such  $\forall (\eta_1, \eta_2) \in (0, \eta], f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2 + \eta_1$ . Thus, for  $p = (v_1(x_1 + x_2) - \alpha, \alpha)$  with  $\alpha \in [v_1(x_2); v_1(x_1 + x_2) - v_1(x_1))$ , we have:

$$\begin{aligned} \|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 &= (p_2 + \eta_1 - p_1)^2 + (p_2 + \eta_1 - p_2)^2 \\ &= (2\alpha - v_1(x_1 + x_2) + \eta_1)^2 + \eta_1^2 \end{aligned} \quad (\text{B.59})$$

Thus,  $\forall p \in I_\alpha^1, g_\Gamma(p) > 0$ .

- Suppose  $p \in I_\alpha^2$  and, thus,  $p = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ . Let's calculate the final prices of  $\Gamma$  if both players play the point price prediction bidding strategy using as initial prediction  $(p_1 - \eta_1, p_2 - \eta_2)$  a slight underestimation of  $p$  with  $\eta_1 > 0, \eta_2 > 0$ . As  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $v_1(x_1 + x_2) - v_1(x_1) \leq \frac{v_2(x_1 + x_2)}{2}$ , player 2 will always bid on both items if their price is below  $v_1(x_1 + x_2) - v_1(x_1)$ . As  $p_1 - \eta_1 < v_1(x_1)$  and  $p_2 - \eta_2 < v_1(x_1 + x_2) - v_1(x_1) < \frac{v_1(x_1 + x_2)}{2} < v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus has been entirely consumed by the rise in price of item 2. This happens when the price of item 2 reaches  $p_2$ . Thus, player 2 acquires item 2 at price  $p_2$ . Then, player 1 will only bid on item 1 until its price reaches her stand-alone value  $p_1 = v_1(x_1)$ . As  $v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$  and player 2 has already acquired item 2, player 2 will also continue to bid on item 1 and, then, acquire it for a final price of  $p_1$ . More formally, for  $p = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ , we have:



$v_1(x_1)$ ),  $\exists \eta > 0$  such that  $\forall (\eta_1, \eta_2) \in (0, \eta]$ ,  $f_\Gamma^1(p_1 - \eta_1, p_2 - \eta_2) = p_1$  and  $f_\Gamma^2(p_1 - \eta_1, p_2 - \eta_2) = p_2$ . Thus, we have:

$$\|f_\Gamma(p_1 - \eta_1, p_2 - \eta_2) - p\|_2^2 = 0 \quad (\text{B.60})$$

Thus,  $g_\Gamma(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) = 0$ .

We have shown that:

- $\forall p \in I_\alpha^1 = \{(v_1(x_1 + x_2) - \alpha, \alpha), \alpha \in [v_1(x_2), v_1(x_1 + x_2) - v_1(x_1)]\}$ ,  $g_\Gamma(p) > 0$ .
- $g_\Gamma(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) = 0$

From these results, the frontier prediction of final prices is easily determined as  $O_\Gamma^{min*} = \{(v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))\}$ .  $\square$

*Proof. Sequence convergence:*

Let's first show that:

$$\begin{aligned} p &\in [0, v_1(x_1)) \times [0, v_1(x_1 + x_2) - v_1(x_1)) \\ &\implies f_\Gamma(p_1, p_2) = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1)) \end{aligned} \quad (\text{B.61})$$

Suppose  $p \in [0, v_1(x_1)) \times [0, v_1(x_1 + x_2) - v_1(x_1))$ . As  $v_1(x_1 + x_2) - v_1(x_1) < v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$ ,  $v_1(x_1 + x_2) - v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_1)$  and  $v_1(x_1 + x_2) - v_1(x_1) \leq \frac{v_2(x_1 + x_2)}{2}$ , player 2 will bid on both items if their price is below  $v_1(x_1 + x_2) - v_1(x_1)$ . As  $p_1 < v_1(x_1)$  and  $p_2 < v_1(x_1 + x_2) - v_1(x_1) < v_1(x_1)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed by item 2. Thus, the price of each item rises symmetrically. Player 1 will stop bidding on item 2 when its price reaches  $v_1(x_1 + x_2) - v_1(x_1)$  and will continue to bid on item 1 until its price reaches  $v_1(x_1)$ . Player 2 will then acquire item 2 for  $v_1(x_1 + x_2) - v_1(x_1)$  and will continue bidding on item 1 until its price reaches  $v_2(x_1 + x_2) - v_2(x_2)$ . As  $v_1(x_1) \leq v_2(x_1 + x_2) - v_2(x_2)$ , player 2 acquires item 1 for  $v_1(x_1)$ . Thus,  $f_\Gamma(p_1, p_2) = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ .

As  $P_0 = (0, 0) < (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$  component-wise and  $P_{t+1} = \frac{1}{t+1}f_\Gamma(P_t) + (1 - \frac{1}{t+1})P_t$ , by a simple induction we have that  $\forall t \in \mathbb{R}_+, P_t \in [0, v_1(x_1)] \times [0, v_1(x_1 + x_2) - v_1(x_1)]$  and, thus,  $\forall t \in \mathbb{R}_+, f_\Gamma(P_t) = (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$ .

The above results directly imply that  $P_t \xrightarrow[t \rightarrow +\infty]{} (v_1(x_1), v_1(x_1 + x_2) - v_1(x_1))$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

**c) Third case:**

$$* v_1(x_2) > \frac{v_1(x_1+x_2)}{2}$$

*Proof. Frontier prediction of final prices:*

It is exactly the symmetric case of the second case of disposition 3. Thus, we have  $O_\Gamma^{min*} = \{(v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))\}$ . □

*Proof. Sequence convergence:*

Let's first show that:

$$\begin{aligned} p &\in [0, v_1(x_1 + x_2) - v_1(x_2)] \times [0, v_1(x_2)] \\ \implies f_\Gamma(p_1, p_2) &= (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2)) \end{aligned} \tag{B.62}$$

Suppose  $p \in [0, v_1(x_1 + x_2) - v_1(x_2)] \times [0, v_1(x_2)]$ . As  $v_1(x_1 + x_2) - v_1(x_2) < v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_1)$ ,  $v_1(x_1 + x_2) - v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_2)$  and  $v_1(x_1 + x_2) - v_1(x_2) \leq \frac{v_2(x_1+x_2)}{2}$ , player 2 will bid on both items if their price is below  $v_1(x_1 + x_2) - v_1(x_2)$ . As  $p_2 < v_1(x_2)$  and  $p_1 < v_1(x_1 + x_2) - v_1(x_2) < v_1(x_2)$ , player 1 will bid on both items until her complementarity surplus is entirely consumed by item 1. Thus, the price of each item rises symmetrically. Player 1 will stop bidding on item 1 when its price reaches  $v_1(x_1 + x_2) - v_1(x_2)$  and will continue to bid on item 2 until its price reaches  $v_1(x_2)$ . Player 2 will then acquire item 1 for  $v_1(x_1 + x_2) - v_1(x_2)$  and will continue bidding on item 2 until its price reaches  $v_2(x_1 + x_2) - v_2(x_1)$ . As  $v_1(x_2) \leq v_2(x_1 + x_2) - v_2(x_1)$ , player 2 acquires item 2 for  $v_1(x_2)$ . Thus,  $f_\Gamma(p_1, p_2) = (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))$ .

As  $P_0 = (0, 0) < (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))$  component-wise and  $P_{t+1} = \frac{1}{t+1}f_\Gamma(P_t) + (1 - \frac{1}{t+1})P_t$ , by a simple induction we have that  $\forall t \in \mathbb{R}_+, P_t \in [0, v_1(x_1 + x_2) - v_1(x_2)] \times [0, v_1(x_2)]$  and, thus,  $\forall t \in \mathbb{R}_+, f_\Gamma(P_t) = (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))$ .

The above results directly imply that  $P_t \xrightarrow[t \rightarrow +\infty]{} (v_1(x_1 + x_2) - v_1(x_2), v_1(x_2))$  and, thus, the sequence  $P_t$  converges to the unique element of  $O_\Gamma^{min*}$ .

□

## Appendix: Example 1 in SAA-c game

Let  $\Gamma$  be the instance of the SAA-c game in *Example 1* (see Section 2.2.3) which is recalled below.

	$v(\{1\})$	$v(\{2\})$	$v(\{1, 2\})$
Agent 1	12	12	12
Agent 2	0	0	20

**Tab. C.1.:** *Example 1*

### C.1 Function $f_\Gamma$

Let  $p = (p_1, p_2)$  be the initial prediction. Let  $f_\Gamma(p) = (f_\Gamma^1(p), f_\Gamma^2(p))$ . We have:

$$\mathbb{E}[f_\Gamma(p)] = \begin{cases} (1, 0) & \text{if } p_1 + p_2 \geq 20, p_1 \leq p_2 \text{ and } p_1 < 12 \quad (R_{green}) \\ (0, 1) & \text{if } p_1 + p_2 \geq 20, p_1 > p_2 \text{ and } p_2 < 12 \quad (R_{purple}) \\ (11, 11.5) & \text{if } p_1 + p_2 < 20 \text{ and } 12 > p_1 > p_2 \quad (R_{red}) \\ (11.5, 11) & \text{if } p_1 + p_2 < 20 \text{ and } p_1 \leq p_2 < 12 \quad (R_{orange}) \\ (11.5, 1) & \text{if } p_1 + p_2 < 20 \text{ and } p_2 \geq 12 \quad (R_{cyan}) \\ (1, 11.5) & \text{if } p_1 + p_2 < 20 \text{ and } p_1 \geq 12 \quad (R_{yellow}) \\ (0, 0) & \text{if } p_1 + p_2 \geq 20 \text{ and } p_1 \geq 12 \text{ and } p_2 \geq 12 \quad (R_{magenta}) \end{cases}$$

$\mathbb{E}[f_\Gamma]$  is represented in *Figure C.1*. The arrows correspond to the direction that the sequence  $p_{t+1} = \frac{1}{t+1} \mathbb{E}[f_\Gamma(p_t)] + (1 - \frac{1}{t+1})p_t$  will take in each coloured region. Each coloured region in *Figure C.1* will be referred to as  $R_{colour}$ . For example,  $R_{red} = \{(p_1, p_2) | p_1 + p_2 < 20, p_2 < p_1 < 12\}$ .

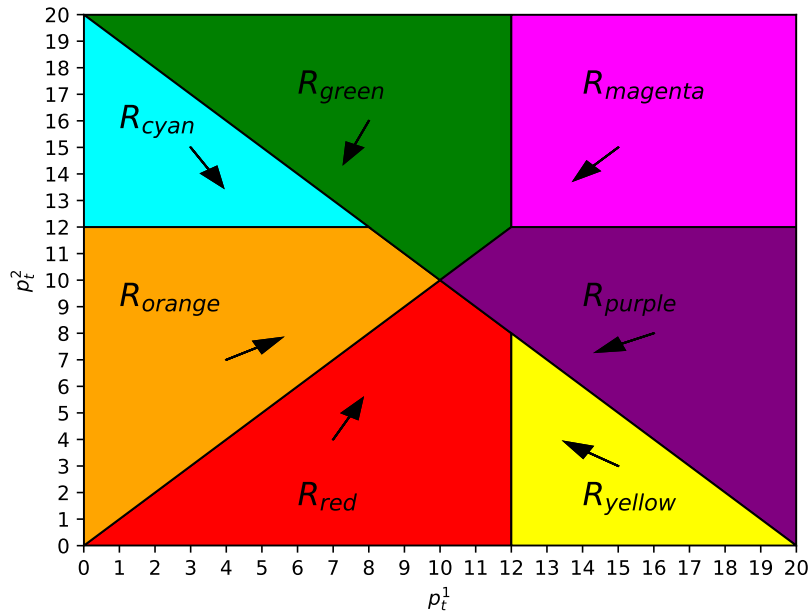


Fig. C.1.: Spatial representation of  $f_{\Gamma}$  in the SAA-c game of Example 1.

## C.2 Convergence of sequence $p_t$ in the SAA-c game of Example 1

Let's show that the sequence  $p_{t+1} = \frac{1}{t+1} \mathbb{E}[f_{\Gamma}(p_t)] + (1 - \frac{1}{t+1})p_t$  with  $P_0 = (0, 0)$  converges to  $p^* = \{(10, 10)\}$ .

The proof is divided into three parts with each a geometrical interpretation:

1. The sequence  $p_t$  is never in  $R_{magenta}$ ,  $R_{cyan}$  and  $R_{yellow}$ .
2. The sequence  $p_t^1 - p_t^2$  converges to zero.
3. The sequence  $p_t^1 + p_t^2$  converges to 20 and conclude that  $p_t \xrightarrow[t \rightarrow +\infty]{} (10, 10)$ .

### C.2.1 $p_t$ is never in $R_{magenta}$ , $R_{cyan}$ and $R_{yellow}$

*Proof.* Notice that  $\max_{p \in \mathbf{R}^{+2}} \mathbb{E}[f_{\Gamma}^1(p)] = \max_{p \in \mathbf{R}^{+2}} \mathbb{E}[f_{\Gamma}^2(p)] = 11.5$ . Hence,  $\forall t \in \mathbb{N}$ ,  $p_t^1 \leq 11.5$  and  $p_t^2 \leq 11.5$ . Thus, the sequence  $p_t$  is never in  $R_{magenta}$ ,  $R_{cyan}$  and  $R_{yellow}$ .

□

### C.2.2 Convergence of $p_t^1 - p_t^2$ to zero

*Proof.* Let's prove that  $U_t = p_t^1 - p_t^2$  converges to zero.

According to the previous result,  $\forall t \in \mathbb{N}$ ,  $p_t^1 \leq 11.5$  and  $p_t^2 \leq 11.5$ . This means that  $p_t$  is never in  $R_{magenta}$ ,  $R_{cyan}$  and  $R_{yellow}$ . Thus,  $p_t$  is only in  $R_{orange}$ ,  $R_{red}$ ,  $R_{purple}$  and  $R_{green}$ . Therefore,

- If  $p_t^1 > p_t^2$ ,  $\mathbb{E}[f_{\Gamma}^1(p_t)] + 0.5 \leq \mathbb{E}[f_{\Gamma}^2(p_t)] \leq \mathbb{E}[f_{\Gamma}^1(p_t)] + 1$
- If  $p_t^1 \leq p_t^2$ ,  $\mathbb{E}[f_{\Gamma}^2(p_t)] + 0.5 \leq \mathbb{E}[f_{\Gamma}^1(p_t)] \leq \mathbb{E}[f_{\Gamma}^2(p_t)] + 1$

We define the sequence  $U_t$  as  $U_t = p_t^1 - p_t^2$  with  $U_0 = 0$ . Thus,  $U_{t+1} = \frac{t}{t+1}U_t + \frac{1}{t+1}(\mathbb{E}[f_{\Gamma}^1(p_t)] - \mathbb{E}[f_{\Gamma}^2(p_t)])$ . Therefore, we have:

- if  $U_t > 0$ :  $\frac{t}{t+1}U_t - \frac{1}{t+1} \leq U_{t+1} \leq \frac{t}{t+1}U_t - \frac{0.5}{t+1}$
- if  $U_t \leq 0$ :  $\frac{t}{t+1}U_t + \frac{0.5}{t+1} \leq U_{t+1} \leq \frac{t}{t+1}U_t + \frac{1}{t+1}$

Let's show by induction that  $\forall t \in \mathbb{N}$ ,  $U_t \in [-\frac{1}{t}, \frac{1}{t}]$ .

1.  $U_1 = 0.5$  so  $U_1 \in [-1, 1]$
2. Suppose  $U_t \in [-\frac{1}{t}, \frac{1}{t}]$ .

If  $U_t > 0$ , then  $U_t \in ]0, \frac{1}{t}]$  and so, following the inequalities on  $U_{t+1}$ , we have:

$$-\frac{1}{t+1} \leq U_{t+1} \leq \frac{1}{t+1} - \frac{0.5}{t+1} \leq \frac{1}{t+1} \quad (\text{C.1})$$

Thus, if  $U_t > 0$ , then  $U_{t+1} \in [-\frac{1}{t+1}, \frac{1}{t+1}]$

If  $U_t \leq 0$ , then  $U_t \in [-\frac{1}{t}, 0]$  and so, following the inequalities on  $U_{t+1}$ , we have:

$$-\frac{1}{t+1} \leq -\frac{1}{t+1} + \frac{0.5}{t+1} \leq U_{t+1} \leq \frac{1}{t+1} \quad (\text{C.2})$$

Thus, if  $U_t \leq 0$ , then  $U_{t+1} \in [-\frac{1}{t+1}, \frac{1}{t+1}]$

Therefore, this proves by induction that  $\forall t \geq 1, U_t \in [-\frac{1}{t}, \frac{1}{t}]$ .

From this result, it trivially follows that  $U_t \xrightarrow[t \rightarrow +\infty]{} 0$ .

□

From now on, we will write  $p_t^2 = p_t^1 + \eta_t$  with  $\eta_t \xrightarrow[t \rightarrow +\infty]{} 0$ . In this part, we have proved that the sequence converges to the diagonal  $p_1 = p_2$ . Thus, given the fact that  $\forall t \in \mathbb{N}$ ,  $p_t$  is either in  $R_{orange}, R_{red}, R_{purple}$  and  $R_{green}$ , we have  $\forall t \geq 1$ :

$$U_{t+1} = \begin{cases} \frac{tU_t}{t+1} - \frac{1}{t+1} & \text{if } U_t > 0 \quad (p_t \in \{R_{red}, R_{purple}\}) \\ \frac{tU_t}{t+1} + \frac{1}{t+1} & \text{if } U_t \leq 0 \quad (p_t \in \{R_{orange}, R_{green}\}) \end{cases} \quad (\text{C.3})$$

As  $\forall t \geq 1, U_t \in [-\frac{1}{t}, \frac{1}{t}]$ ,  $\eta_t \in [-\frac{1}{t}, \frac{1}{t}]$ .

### C.2.3 Convergence of $p_t$ to (10,10) and geometrical interpretation

*Proof.*  $\forall t, p_t \in \{R_{green}, R_{purple}, R_{red}, R_{orange}\}$ . We define  $W_t = p_t^1 + p_t^1 - 20$ . We have:

$$\begin{aligned} W_{t+1} &= p_{t+1}^1 + p_{t+1}^2 - 20 \\ &= \frac{\mathbb{E}[f_{\Gamma}^1(p_t)] + \mathbb{E}[f_{\Gamma}^2(p_t)]}{t+1} + \frac{t(p_t^1 + p_t^2)}{t+1} - 20 \\ &= \frac{\mathbb{E}[f_{\Gamma}^1(p_t)] + \mathbb{E}[f_{\Gamma}^2(p_t)]}{t+1} + \frac{tW_t}{t+1} + \frac{20t}{t+1} - 20 \\ &= \frac{\mathbb{E}[f_{\Gamma}^1(p_t)] + \mathbb{E}[f_{\Gamma}^2(p_t)]}{t+1} + \frac{tW_t}{t+1} - \frac{20}{t+1} \end{aligned} \quad (\text{C.4})$$

Thus,

$$W_{t+1} = \begin{cases} \frac{tW_t}{t+1} + \frac{5}{2(t+1)} & \text{if } W_t < 0 \quad (p_t \in \{R_{red}, R_{orange}\}) \\ \frac{tW_t}{t+1} - \frac{19}{t+1} & \text{if } W_t \geq 0 \quad (p_t \in \{R_{purple}, R_{green}\}) \end{cases} \quad (\text{C.5})$$

Similarly to the preceding proof, we can show by induction that  $\forall t \geq 1, W_t \in [-\frac{19}{t}, \frac{5}{2t}]$  as  $W_1 = \frac{5}{2}$ . **Hence,**  $W_t \xrightarrow{t \rightarrow +\infty} 0$ . **Thus,**  $p_t^1 \xrightarrow{t \rightarrow +\infty} 10$  **and**  $p_t^2 \xrightarrow{t \rightarrow +\infty} 10$ .

Now, let's show that  $p_t$  belongs to the diamond defined by the points  $(10 - \frac{10}{t}, 10 - \frac{9}{t})$ ,  $(10 - \frac{9}{t}, 10 - \frac{10}{t})$ ,  $(10 + \frac{7}{4t}, 10 + \frac{3}{4t})$  and  $(10 + \frac{3}{4t}, 10 + \frac{7}{4t})$  when  $t \geq 1$ .

$\forall t \geq 1$ , we have:

$$\begin{aligned} -\frac{19}{t} &\leq W_t \leq \frac{5}{2t} \\ \Leftrightarrow -\frac{19}{t} &\leq 2p_t^1 + \eta_t - 20 \leq \frac{5}{2t} \\ \Leftrightarrow -\frac{19}{2t} + 10 - \frac{\eta_t}{2} &\leq p_t^1 \leq \frac{5}{4t} + 10 - \frac{\eta_t}{2} \end{aligned} \quad (\text{C.6})$$

As  $\eta_t \in [-\frac{1}{t}, \frac{1}{t}]$ , we have that:

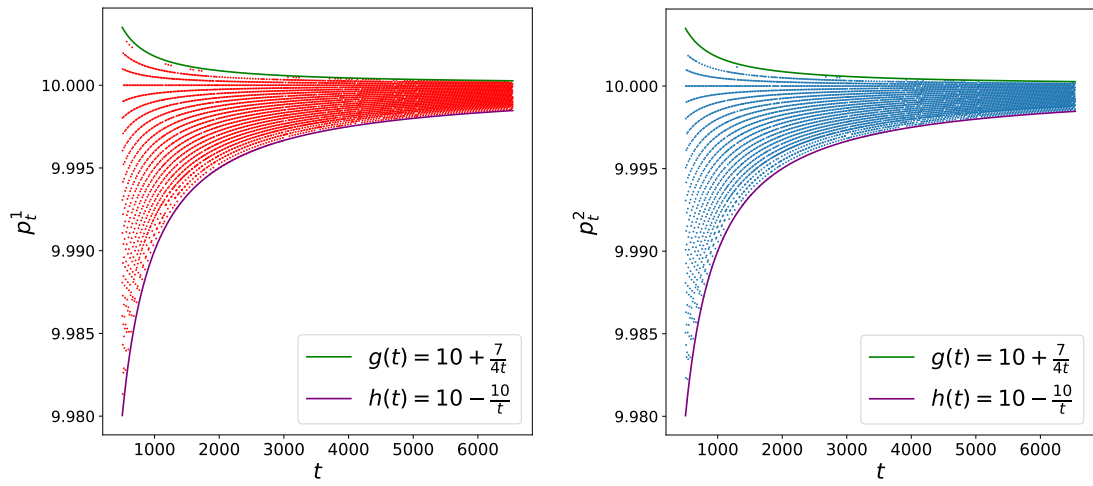
$$\forall t \geq 1, 10 - \frac{10}{t} \leq p_t^1 \leq 10 + \frac{7}{4t} \quad \text{and} \quad 10 - \frac{10}{t} \leq p_t^2 \leq 10 + \frac{7}{4t} \quad (\text{C.7})$$

The sequence  $p_t^1$  and  $p_t^2$  along with their upper-bound  $g(t) = 10 + \frac{7}{4t}$  and lower-bound  $h(t) = 10 - \frac{10}{t}$  are illustrated in Figure C.2.

Geometrically, we have first shown, through  $U_t \in [-\frac{1}{t}, \frac{1}{t}]$ , that  $p_t$  was between the line  $p_1 - p_2 = -\frac{1}{t}$  and  $p_1 - p_2 = \frac{1}{t}$  which are parallel to the line  $p_1 = p_2$ . Then, in this last part, through  $W_t \in [-\frac{19}{t}, \frac{5}{2t}]$ , we showed that  $p_t$  is also between the lines  $p_1 + p_2 = 20 - \frac{19}{t}$  and  $p_1 + p_2 = 20 + \frac{5}{2t}$  which are parallel to the line  $p_1 + p_2 = 20$ . Using both of these information, we know that  $p_t$  belongs to the diamond defined by the points  $(10 - \frac{10}{t}, 10 - \frac{9}{t})$ ,  $(10 - \frac{9}{t}, 10 - \frac{10}{t})$ ,  $(10 + \frac{7}{4t}, 10 + \frac{3}{4t})$  and  $(10 + \frac{3}{4t}, 10 + \frac{7}{4t})$  which progressively converges to the point  $(10, 10)$ .

□





**Fig. C.2.:** Convergence rate of  $p_t^1$  and  $p_t^2$  in the SAA-c game of *Example 1*.

**Titre :** Enchérir efficacement dans l'enchère ascendante simultanée en utilisant la recherche arborescente Monte Carlo

**Mots clés :** Recherche arborescente Monte Carlo, Enchère ascendante simultanée, information incomplète, aversion au risque, théorie des jeux

**Résumé :** Depuis son introduction en 1994 aux États-Unis, l'enchère ascendante simultanée (SAA) est devenue le mécanisme privilégié pour les enchères du spectre licencié. Avec des investissements dépassant parfois le milliard d'euros, une stratégie d'enchérissement performante devient cruciale pour les opérateurs mobiles.

Malgré son importance, il existe un manque de recherche dédiée à la création d'une stratégie d'enchérissement performante dans le cadre du SAA. La complexité intrinsèque du jeu associé à l'enchère SAA rend son analyse ardue pour la théorie des enchères et les méthodes exactes de résolution de jeux. De plus, ce mécanisme engendre des problèmes stratégiques tels que le problème d'exposition, ajoutant une couche de complexité supplémentaire à son étude.

Cette thèse propose l'utilisation de la recherche arborescente de Monte Carlo (MCTS) pour calculer une stratégie d'enchérissement performante au sein du SAA. Les six chapitres de la thèse sont structurés comme suit. Le premier chapitre présente les mécanismes d'enchères du spectre licencié, soulignant leurs avantages et inconvénients.

Le deuxième chapitre détaille le problème spécifique de l'enchérisseur dans le SAA, ainsi que certains travaux connexes. Le troisième chapitre propose une synthèse concise des méthodes traditionnelles de recherche dans les jeux avec des adversaires, en mettant particulièrement l'accent sur le MCTS. Les chapitres quatre à six sont dédiés à la création de l'algorithme MCTS pour calculer une stratégie performante. Le quatrième chapitre modélise l'enchère SAA comme un jeu au tour par tour à N-joueurs à information parfaite et complète, avec des expériences numériques sur des instances de petite taille. Le cinquième chapitre modélise l'enchère comme un jeu simultané à N-joueurs à information complète, avec des contraintes budgétaires et d'éligibilité, et les résultats sont obtenus sur des instances de taille réelle. Le sixième chapitre considère le jeu à information incomplète pour modéliser les incertitudes de la réalité. Pour chaque modèle, un algorithme surpassant largement ceux de la littérature est proposé, traitant notamment le problème d'exposition. De plus, une méthode de prédiction des prix finaux est développée tout au long des chapitres, sur laquelle chaque algorithme MCTS s'appuie.

**Title :** Bidding efficiently in Simultaneous Ascending Auctions using Monte Carlo Tree Search

**Keywords :** Monte Carlo Tree Search, Simultaneous Ascending Auction, incomplete information, risk-aversion, game theory

**Abstract :** Since its introduction in 1994 in the United States, the Simultaneous Ascending Auction (SAA) has become the privileged mechanism for spectrum auctions. As sometimes billions of euros are at stake in an SAA, and a mobile operator's business plan highly depends on the auction outcome, establishing an efficient bidding strategy is crucial.

Despite the importance of this problem, there is a lack of research dedicated to developing an efficient bidding strategy for the SAA. The intrinsic complexity of the SAA makes its analysis very challenging for auction theory and exact game resolution methods. Additionally, the mechanism introduces strategic issues such as the exposure problem, adding an extra layer of complexity to its study.

This thesis proposes the use of Monte Carlo Tree Search (MCTS) to compute an efficient bidding strategy for the SAA. The six chapters of the thesis are structured as follows. The first chapter introduces spectrum auction mechanisms, highlighting their pros

and cons. The second chapter details the bidding problem in the SAA, along with relevant related research. The third chapter provides a summary of adversarial search methods, with a specific focus on MCTS. Chapters four to six are dedicated to developing an efficient MCTS bidding strategy for the SAA. The fourth chapter considers a turn-based deterministic model of the SAA with perfect and complete information. Numerical experiments are only undertaken on small instances. The fifth chapter considers an n-player simultaneous move model of SAA with incomplete information. Extensive numerical experiments are undertaken on instances of realistic size. The sixth chapter extends the preceding game to incomplete information to introduce uncertainties. For each model, an algorithm that significantly outperforms state-of-the-art bidding strategies is proposed, notably by better tackling the exposure problem. Moreover, a final price prediction method is developed throughout the chapters, upon which each MCTS algorithm relies.