UNIVERSITÉ DE LORRAINE | BIBLIOTHÈQUES UNIVERSITAIRES

AVERTISSEMENT
DROIT D'AUTEUR

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document. Toute contrefaçon, plagiat ou reproduction illicite encourt une poursuite pénale.

Contact bibliothèque : ddoc-theses-contact@univ-lorraine.fr
(Cette adresse ne permet pas de contacter les auteurs)

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

UNIVERSITÉ DE LORRAINE
34 COURS LÉOPOLD — CS 25233
54052 NANCY CEDEX
TÉL : 03 54 50 54 00
CONTACT@UNIV-LORRAINE.FR
WWW.UNIV-LORRAINE.FR

UNIVERSITÉ DE LORRAINE

École doctorale IAEM Lorraine

Généralisations du jeu de Blotto : du différentiel stochastique au jeu à population infinie

THÈSE

présentée et soutenue publiquement le 22 avril 2026

pour l'obtention du

Doctorat de l'Université de Lorraine

(mention mathématiques)

par

Mohamed-Amine Hazzami

Composition du jury

Président du jury et Rapporteur :	Thibaut Mastrolia	Assistant Professor à UC Berkeley
Rapporteuse :	Caroline Hillairet	Professeure à ENSAE-Paris, CREST
Examinatrices et Examinateurs :	Pauline Barrieu	Professeure à London School of Economics
	Marie Laclau	Associate Professor à HEC Paris, CNRS
	Jérôme Lohéac	Chargé de Recherche au CRAN, CNRS
	Pascal Moyal	Professeur à l'Université de Lorraine
Directeur de thèse :	Nabil Kazi-Tani	Professeur à l'Université de Lorraine
Codirecteur de thèse :	Vineeth Satheeskumar-Varma	Chargé de Recherche au CRAN, CNRS

Institut Élie Cartan de Lorraine et CNRS — UMR 7502

Mis en page avec la classe thesul.

Remerciements

Je tiens d'abord à remercier celles et ceux sans qui ce doctorat n'aurait pas eu la même trajectoire.

Je suis profondément reconnaissant envers mon directeur de thèse, Nabil Kazi-Tani, et mon codirecteur, Vineeth Satheeskumar-Varma. J'ai eu une chance rare : être encadré par des personnes qui, bien au-delà de leur rôle, ont été pour moi comme des grands frères. Leur disponibilité, leur aide, leur présence, leur solidité dans les moments où l'on doute, et leur exigence quand il fallait avancer m'ont porté tout au long du chemin. J'ai énormément appris à leurs côtés. Merci pour la confiance, pour la patience, pour les conseils toujours utiles, et pour ce sentiment très précieux de ne jamais être seul face aux difficultés.

Je remercie également les membres de mon jury, en particulier Caroline Hillairet et Thibaut Mastrolia, qui ont accepté d'être rapporteurs de cette thèse et dont le travail de lecture et de rapport représente un investissement considérable. Je remercie aussi Pauline Barrieu, Marie Laclau, Jérôme Lohéac et Pascal Moyal. Je leur suis très reconnaissant pour la lecture attentive de ce manuscrit, pour leurs remarques et pour les discussions scientifiques auxquelles ce travail a donné lieu.

Je tiens aussi à remercier Pascal Moyal, Antoine Lejay et Antoine Henrot pour leur guidance au moment de commencer une thèse. Dans les premières étapes, leurs conseils ont permis de rendre le départ moins intimidant.

Je veux ensuite remercier mes amis du Discord : Mehdi, Mehdi et Mehdi (dans cet ordre précis), Walid et Walid (là encore), Omar, Yahya, Salah, Ghassane, Hamza, Anas, Othmane, Aymane, Soufiane, Ismail, Souhail et tous ceux que j'ai oubliés. Même quand on est loin, même quand les rythmes ne se synchronisent pas, vous avez été là. Merci pour les messages, les appels, les silences qui ne gênent pas, et surtout pour ces séances de débats, de jeux, et de bêtises qui permettent de se déconnecter.

Je veux aussi remercier mes amis doctorants de l'IECL : Simon, Maxime, Kai, Aurélie, Nicolas et Ruben. Vous avez formé un groupe chaleureux et bienveillant, avec qui on partage les petites victoires, les périodes denses, et parfois aussi les journées où rien n'avance. Merci pour l'ambiance, la solidarité, les discussions (scientifiques ou pas), les sorties, les voyages, et ce sentiment d'appartenance qui change tout.

Je veux adresser un merci particulier à Virginie. Ta bienveillance, ton aide, ton écoute et ton amitié ont compté bien plus que tu ne l'imagines. Tu as eu cette manière rare de rendre les choses plus légères, sans jamais nier leur difficulté. Dans un doctorat, il y a des moments où une présence sincère fait toute la différence : merci d'avoir été cette présence, si souvent.

Enfin, je remercie mon père, ma mère, mes sœurs et le reste de ma famille. Pour l'amour, la patience, la fierté, les encouragements, et pour avoir été là, même quand je n'avais pas grand-chose à raconter, même quand j'étais fatigué, même quand le doctorat débordait un peu sur tout le reste. Ce manuscrit porte votre soutien.

Pour ma famille.

Table des matières

Chapitre 1
Introduction **1**

1.1 Historique et contexte actuel 2
1.2 Le jeu de Blotto 6
1.2.1 Cadre général 6
1.2.2 Règle de gain 7
1.2.3 Quelques variations du problème 7
1.2.4 Stratégies mixtes 8
1.3 Résolution du jeu de Blotto 9
1.3.1 Premiers résultats de Borel 9
1.3.2 Modèle continu de Gross et Wagner 9
1.3.3 Avancées sur le jeu de Blotto 12
1.4 Jeux différentiels stochastiques et extension du jeu de Blotto 14
1.4.1 Formulation des jeux différentiels stochastiques 14
1.4.2 Les gains des joueurs 15
1.4.3 Jeux différentiels stochastiques markoviens 16
1.4.4 Approche par les EDP 16
1.4.5 Jeux à somme nulle 18
1.4.6 Contribution : jeu de Blotto différentiel stochastique symétrique 20
1.4.7 EDSR et jeux différentiels stochastiques 26
1.4.8 Contribution : jeu de Blotto différentiel stochastique asymétrique 28
1.4.9 Numérique 29
1.5 Théorie des contrats et jeux du Principal-Agent 35
1.5.1 Jeux de Principal-Agent différentiels stochastiques 35
1.5.2 Contribution : Jeu P-A différentiel stochastique de Blotto pessimiste . . . 42
1.5.3 Dérivation de l'EDP HJBI 50
1.6 Contribution : Une perspective à population infinie 59
1.6.1 Introduction du problème de champ moyen du Colonel Blotto 59

1.6.2 Un manque de structure 60
1.6.3 Un modèle-jouet de champ moyen 62
1.6.4 Le jeu d'allocation à champ moyen avec pénalisation entropique 62
1.6.5 Construction d'un équilibre à champ moyen du jeu de Blotto 64

Chapitre 2
Stochastic differential Blotto game

2.1 Introduction 71
2.2 Model and Assumptions 73
2.3 The symmetric case 74
2.4 Particular cases in the symmetric setting 77
2.4.1 Fixed payoff for each battlefield 77
2.4.2 Elections inspired payoff 78
2.4.3 Winning all the battlefields 79
2.4.4 Spectral games on graphs 79
2.5 The asymmetric case 82
2.5.1 Characterization of a Nash Equilibrium 82
2.5.2 Numerics 84

Chapitre 3
Principal-Agent Blotto game

3.1 Introduction 89
3.2 Model and Assumptions 91
3.3 Solving the game between the agent and the hacker 93
3.4 Collapsing the Principal Hacker's problem 95
3.5 Solving the Principal's problem 98
3.5.1 Deriving the HJB PDE 98
3.5.2 Pointwise optimization in the Hamiltonian 104
3.5.3 Resulting PDE and BSDE representation 110
3.6 Numerics for the Principal–Agent game 111

Chapitre 4
The Mean-Field Blotto game

4.1 Introduction 113
4.2 Model and assumptions 114
4.3 A lack of structure 116
4.4 A toy mean-field model 118

4.5 Entropy-regularized mean-field allocation game 118
4.5.1 Entropy-regularized best response 120
4.5.2 General mean-field equilibrium 120
4.6 Constructing a mean-field equilibrium for the Blotto game 121
4.6.1 Homogeneous case and exchangeable equilibria 121
4.6.2 Vector thresholds and weighted battlefields 122
4.6.3 Explicit equilibria in low dimensions 123

Annexes **127**

Annexe A
Preuves **127**

Annexe B
Modèle-jouet à champ moyen **131**

Bibliographie **135**

Abstract **141**

Résumé **142**

Chapitre 1

Introduction

Sommaire

1.1 Historique et contexte actuel 2
1.2 Le jeu de Blotto 6
1.2.1 Cadre général 6
1.2.2 Règle de gain 7
1.2.3 Quelques variations du problème 7
1.2.4 Stratégies mixtes 8
1.3 Résolution du jeu de Blotto 9
1.3.1 Premiers résultats de Borel 9
1.3.2 Modèle continu de Gross et Wagner 9
1.3.3 Avancées sur le jeu de Blotto 12
1.4 Jeux différentiels stochastiques et extension du jeu de Blotto 14
1.4.1 Formulation des jeux différentiels stochastiques 14
1.4.2 Les gains des joueurs 15
1.4.3 Jeux différentiels stochastiques markoviens 16
1.4.4 Approche par les EDP 16
1.4.5 Jeux à somme nulle 18
1.4.6 Contribution : jeu de Blotto différentiel stochastique symétrique 20
1.4.7 EDSR et jeux différentiels stochastiques 26
1.4.8 Contribution : jeu de Blotto différentiel stochastique asymétrique 28
1.4.9 Numérique 29
1.5 Théorie des contrats et jeux du Principal-Agent 35
1.5.1 Jeux de Principal-Agent différentiels stochastiques 35
1.5.2 Contribution : Jeu P-A différentiel stochastique de Blotto pessimiste . . 42
1.5.3 Dérivation de l'EDP HJBI 50
1.6 Contribution : Une perspective à population infinie 59
1.6.1 Introduction du problème de champ moyen du Colonel Blotto 59
1.6.2 Un manque de structure 60
1.6.3 Un modèle-jouet de champ moyen 62
1.6.4 Le jeu d'allocation à champ moyen avec pénalisation entropique 62
1.6.5 Construction d'un équilibre à champ moyen du jeu de Blotto 64

1.1 Historique et contexte actuel

L'allocation de ressources constitue un problème fondamental en optimisation, mais plus largement encore dans l'ensemble des sciences appliquées et des systèmes décisionnels. Il s'agit de déterminer comment répartir au mieux un budget limité entre plusieurs actions possibles, selon des critères d'efficacité, de survie ou de performance. Et c'est dans ce cadre que s'écrivent les contributions de cette thèse.

Ce questionnement n'est pas exclusivement humain. En écologie, par exemple, Owen-Smith [62] s'intéresse aux liens entre comportement alimentaire et dynamique de population, et étudie, au moyen d'un modèle dynamique dépendant de l'état, la façon dont de jeunes animaux devraient répartir la biomasse gagnée entre croissance corporelle et stockage de réserves dans un environnement saisonnier et stochastique, en particulier pour de grands herbivores. Il montre notamment que les animaux devraient stocker le moins de réserves possible, aussi tard que possible, avec un niveau minimal de stockage qui augmente lorsque l'incertitude environnementale est plus forte. Plus généralement, de nombreux travaux en biologie et en écologie montrent que la dynamique de croissance, de stockage et de reproduction peut être décrite comme un problème d'allocation de ressources sous contrainte.

Chez l'humain, des formes d'arbitrage stratégique entre usages concurrents des ressources apparaissent dès l'Antiquité. Dans le domaine militaire, les traités de stratégie, comme *L'Art de la guerre* de Sun Tzu [78], insistent sur la manière de répartir et de déplacer les forces sur plusieurs fronts afin d'atteindre un objectif global. Le jeu de Blotto, qu'on introduira et traitera tout au long de cette thèse, se formule dans le même esprit. De façon analogue, les textes anciens sur l'administration de la cité discutent déjà de la répartition de ressources limitées (financières, humaines ou matérielles) entre différents objectifs politiques ou économiques.

Bien plus tard, à la fin du XIXe siècle, la question de l'allocation optimale des ressources est reprise et formalisée par plusieurs économistes. Ces auteurs cherchent à décrire, dans un cadre analytique, la façon dont un agent rationnel répartit ses ressources en fonction des gains attendus.

Par exemple, Jevons [47] développe l'idée que la valeur économique d'un bien doit être reliée à son utilité marginale, et propose de traiter l'économie comme une science mathématique fondée sur des fonctions d'utilité et des contraintes de rareté. Son approche contribue à installer l'idée que les décisions économiques peuvent être modélisées comme des problèmes d'optimisation sous contrainte.

Walras [81] formalise un équilibre général : un système d'équations relie l'offre, la demande et les prix sur l'ensemble des marchés simultanément. Dans ce cadre, les agents choisissent des allocations de biens en fonction des prix, et ces décisions sont interdépendantes via un système global de prix. La question de l'existence d'un équilibre devient alors un problème mathématique bien posé, qui donnera lieu à une vaste littérature d'analyse économique rigoureuse.

Pareto [64] introduit la notion d'efficacité (ou d'optimalité) au sens de Pareto pour caractériser des situations où il n'est plus possible d'améliorer la situation d'un agent sans détériorer celle d'un autre. Cette notion fournit un critère d'efficacité pour comparer des allocations de ressources du point de vue collectif, sans imposer de mesure unique du bien-être.

Ces travaux font partie des premières tentatives systématiques de modéliser des choix économiques rationnels sous contrainte de ressources. Toutefois, ils restent centrés sur des décisions individuelles (ou agrégées), et ne traitent pas encore explicitement la dimension stratégique de l'interaction entre plusieurs agents aux intérêts potentiellement divergents. Ce n'est qu'avec l'émergence de la théorie des jeux, au début du XXe siècle, que cette interaction devient un objet d'étude central.

Effectivement, au début du XXe siècle, la modélisation des situations de décision évolue vers une prise en compte explicite de l'interaction stratégique entre plusieurs agents. Zermelo [82], dans un article souvent considéré comme l'un des points de départ de la théorie des jeux formelle, étudie des jeux à deux joueurs, sans hasard et avec un nombre fini de positions, en prenant le jeu d'échecs comme exemple. La littérature moderne résume ce résultat sous le nom de théorème de Zermelo : dans tout jeu fini à deux joueurs, à information parfaite et sans chance, soit un joueur peut forcer la victoire, soit les deux peuvent au moins garantir un résultat de non-défaite [72].

Quelques années plus tard, Borel [12] propose une première étude systématique des jeux à deux joueurs à somme nulle et discute l'usage de méthodes analytiques (en particulier les équations intégrales) pour en étudier les stratégies. Ses travaux ouvrent la voie à l'idée de traiter des situations de conflit comme des objets mathématiques bien définis, et introduisent plusieurs exemples de jeux, dont le jeu de Blotto, qui seront repris et généralisés par la suite, y compris dans cette thèse.

La formalisation d'ensemble de la théorie des jeux est proposée dans l'ouvrage de von Neumann et Morgenstern [61]. Cet ouvrage systématise l'analyse des jeux à somme nulle en s'appuyant notamment sur le théorème du minimax de von Neumann, et propose une représentation générale des interactions économiques sous forme de jeux de stratégie. Il fournit aussi un cadre pour relier comportements stratégiques et fonctions d'utilité, en particulier via l'axiomatisation de l'utilité espérée.

L'ensemble de ces contributions établit les fondements d'une modélisation stratégique de l'allocation de ressources entre plusieurs agents, qui sera ultérieurement appliquée à des domaines variés tels que l'économie, la cyber-sécurité ou les sciences politiques.

Depuis, le jeu de Blotto a connu un développement significatif et a été appliqué à une grande variété de contextes stratégiques contemporains. Basé sur la répartition compétitive de ressources limitées sur plusieurs fronts, il fournit un cadre générique pour modéliser une large classe de problèmes d'optimisation en contexte conflictuel. On peut citer, parmi d'autres, les exemples suivants.

En économie et en science politique, les jeux de Colonel Blotto servent à analyser des situations d'allocation concurrentielle de ressources entre plusieurs « champs de bataille ». Kovenock et Roberson [51] étudient une variante coalitionnelle du jeu de Colonel Blotto dans laquelle deux joueurs peuvent former une alliance avant de participer séparément à deux jeux de Colonel Blotto disjoints contre un adversaire commun. Ils montrent que des transferts unilatéraux peuvent apparaître au sein de l'alliance, modifier la stratégie de l'adversaire et bénéficier aux alliés, avec une application explicite à l'économie des alliances.

En cybersécurité, Ferdowsi et al. [28] proposent un cadre de jeu de Colonel Blotto pour analyser l'allocation de ressources de sécurité dans des systèmes cyber-physiques interconnectés au

sein d'une ville intelligente. Ils formulent l'interaction entre un attaquant et un défenseur comme un jeu de Colonel Blotto dans lequel chacun répartit un budget limité entre plusieurs composantes du système, et comparent, à l'équilibre du jeu, les gains des joueurs selon que l'attaquant connaît ou non la structure des interdépendances entre ces composantes. Dans un autre travail, Ferdowsi et al. [29] s'intéressent à la protection des capteurs d'infrastructures critiques interdépendantes (telles que les réseaux d'électricité, de gaz ou d'eau) et proposent un cadre de jeu de Colonel Blotto pour capturer l'interdépendance entre les actions de l'attaquant et de l'administrateur. Ils dérivent l'équilibre en stratégies mixtes et montrent que l'administrateur peut randomiser de manière stratégique entre différents niveaux de protection des capteurs afin de tromper l'attaquant.

En marketing, De Silva et al. [23] considèrent un duopole dans lequel deux entreprises offrent un service ou un produit à des clients répartis sur plusieurs régions géographiques (par exemple des pays, provinces ou États). Chaque région est caractérisée par des paramètres de propagation et de restauration du service, et le taux de propagation peut être contrôlé par un investissement local. Les auteurs formulent ce problème comme un jeu de Stackelberg dans lequel les entreprises choisissent une stratégie d'allocation de leur budget publicitaire entre les différentes régions de manière à maximiser leur part de marché. En caractérisant les équilibres de Stackelberg (optimistes et pessimistes) du jeu, ils décrivent les configurations dans lesquelles la solution prend la forme d'un scénario où le gagnant prend tout. Les jeux de Stackelberg seront introduits plus tard dans ce manuscrit.

Enfin, dans la littérature de théorie des jeux évolutionnaires, Hernández et Zanette [41] proposent une formulation évolutionnaire du jeu de Colonel Blotto en tant que modèle d'allocation de ressources. Dans ce jeu, deux joueurs de dotations totales différentes doivent répartir entièrement leurs ressources entre un ensemble d'items, chaque item étant remporté par le joueur qui y alloue la plus grande quantité de ressources, et le gain correspondant au nombre d'items gagnés. Les auteurs montrent qu'une formulation évolutionnaire, dans laquelle une population de joueurs adapte ses stratégies par imitation et fluctuations aléatoires, permet d'interpréter les stratégies optimales comme des états d'équilibre d'un processus dynamique et de décrire comment des joueurs faiblement dotés peuvent, dans certaines conditions, maximiser leurs gains malgré leur désavantage initial.

Plan de la thèse et résultats importants

Cette thèse est consacrée à l'étude mathématique de jeux d'allocation de ressources, en particulier ceux inspirés du jeu de Blotto. Elle est structurée selon le plan suivant :

On commence par introduire la version formelle du jeu de Blotto : deux joueurs répartissent des ressources limitées sur plusieurs champs d'affrontement, dans un cadre à information parfaite et à somme nulle. Les règles de victoire sont définies champ par champ. On présente les variantes classiques du modèle, et ses formulations élémentaires, et les résultats fondamentaux qui y sont associés. On recense ensuite les principales extensions du modèle de Blotto dans la littérature.

On présente après la modélisation dynamique à travers les jeux différentiels stochastiques. Ces jeux sont associés à des équations de type HJBI, permettant d'analyser les stratégies dans un cadre temporel continu. Le lien entre équilibres stochastiques et solutions d'équations aux dérivées partielles est discuté.

Une première contribution est ensuite proposée, avec une extension différentielle stochastique du jeu de Blotto. Des exemples sur des graphes, où les cibles sont représentées comme des arêtes, sont analysés. Des stratégies optimales explicites sont obtenues dans certains problèmes d'optimisation spécifiques.

Par la suite, on discute de la résolution de ces jeux différentiels stochastiques via des équations différentielles stochastiques rétrogrades (EDSR). On en profite ensuite pour généraliser le jeu de Blotto différentiel stochastique dans un cadre plus large, et le résoudre numériquement.

Un autre type de jeu est ensuite introduit : les jeux hiérarchisés de type Stackelberg ou Principal–Agent, dans lesquels un joueur peut s'engager en premier. On examine leurs formulations dynamiques et stochastiques, ainsi que les nouvelles contraintes qu'ils induisent sur les stratégies et les équilibres.

Une deuxième contribution est présentée dans un cadre de type Principal–Agent, avec incertitude sur les préférences de l'adversaire. Ce cadre est modélisé comme un jeu de Blotto entre deux agents représentant des principaux antagonistes, dans un scénario pessimiste. Un résultat de réduction simplifiant le problème a été prouvé. De plus, des résultats d'existence d'équilibres sont établis, et des stratégies optimales sont construites explicitement, pour l'incitation optimale ainsi que pour les contrôles optimaux.

Une dernière contribution est présentée dans la suite. Il s'agit de l'étude d'un jeu d'allocation de ressources généralisant le jeu de Blotto, dans un cadre de population infinie. On donne pour ce jeu une réponse optimale pour chaque agent, ainsi qu'un résultat d'existence d'équilibre. De plus, on construit explicitement un équilibre pour le cas du jeu de Blotto.

1.2 Le jeu de Blotto

Le jeu de Blotto est un jeu stratégique entre deux joueurs, dans lequel chacun dispose d'un budget limité de ressources qu'il doit répartir entre plusieurs champs d'affrontement. Sur chaque champ, le joueur ayant consacré le plus de ressources remporte ce champ. Le gain global de chaque joueur dépend du nombre et des valeurs des champs remportés.

Sous sa forme classique, ce jeu s'inscrit dans le cadre des jeux à somme nulle, c'est-à-dire que le gain d'un joueur est opposé à celui de son adversaire à une constante près. Il s'agit d'un cas particulier de compétition pure, où les intérêts sont parfaitement antagonistes.

Imaginons, comme dans la figure ci-dessous, deux colonels, chacun disposant d'un certain nombre de soldats, envoyés simultanément sur une série de champs de bataille. Chacun des 5 champs représente un affrontement indépendant, et le colonel qui y envoie le plus de soldats remporte ce champ. Le but est simple : gagner un maximum de champs de bataille.

Cependant, les ressources sont limitées : chaque colonel doit décider comment répartir ses soldats sans connaître les décisions de l'adversaire. Une répartition trop uniforme risque de perdre face à une stratégie plus ciblée ; une concentration excessive sur quelques champs peut laisser trop de terrain à l'adversaire. Ce dilemme stratégique est au cœur du jeu de Blotto.

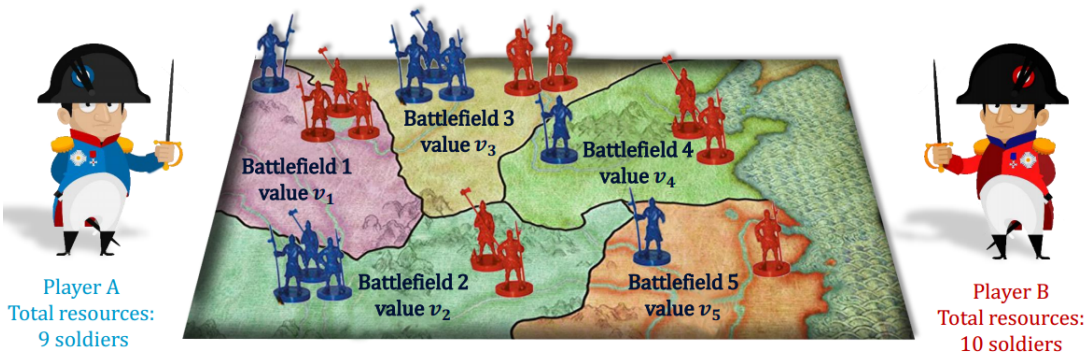


FIGURE 1.1 – Visualisation simplifiée du jeu de Blotto : deux colonels répartissent leurs soldats sur cinq champs de bataille. Source : Site de Dong Quan Vu [79].

Ce cadre, bien qu'inspiré de conflits militaires, sert aujourd'hui de modèle abstrait pour une variété de situations où des ressources limitées doivent être allouées entre des objectifs concurrents : campagnes électorales, défense de systèmes critiques, stratégies publicitaires, ou encore compétition biologique.

1.2.1 Cadre général

Soient $m \in \mathbb{N}^*$ le nombre de champs de bataille (ou cibles), $\alpha > 0$ le budget du joueur A, et $\beta > 0$ celui du joueur B.

Chaque joueur choisit une répartition de son budget sous forme de vecteur :

$$x = (x_1, \ldots, x_m) \in \mathbb{R}^m_{\geq 0} \quad \text{tel que} \quad \sum_{i=1}^{m} x_i = \alpha$$

$$y = (y_1, \ldots, y_m) \in \mathbb{R}^m_{\geq 0} \quad \text{tel que} \quad \sum_{i=1}^{m} y_i = \beta$$

où x_i (resp. y_i) désigne la quantité de ressource allouée par A (resp. B) au champ i.

Chaque champ est associé à une valeur $\lambda_i > 0$ représentant son importance stratégique ou son poids dans la fonction de gain globale.

1.2.2 Règle de gain

Sur chaque champ $i \in \{1, \ldots, m\}$, on attribue un gain, donné par une fonction $g_i : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, définie comme suit :

$$g_i(x_i, y_i) = \begin{cases} \lambda_i & \text{si } x_i > y_i \\ \lambda_i/2 & \text{si } x_i = y_i \\ 0 & \text{si } x_i < y_i \end{cases}$$

Le gain total du joueur A est :

$$G(x, y) = \sum_{i=1}^{m} g_i(x_i, y_i).$$

On remarque que les quantités $g_i(x_i, y_i)$ ne dépendent que de la différence $(x_i - y_i)$.

Le jeu étant à somme nulle, le gain de B est $\sum_{i=1}^{m} \lambda_i - G(x, y)$.

1.2.3 Quelques variations du problème

Version discrète. Les ressources sont représentées par des unités entières, souvent dans un cadre militaire ou électoral, par exemple dans la figure 1.1.

Version continue. Les ressources sont continues et divisibles, souvent utilisée en économie ou en cybersécurité pour modéliser une allocation fluide.

Version symétrique. Les deux joueurs disposent du même budget ($\alpha = \beta$), ce qui simplifie l'analyse stratégique, plutôt utilisée pour les compétitions équilibrées.

Version asymétrique. Les joueurs ont des budgets différents ($\alpha \neq \beta$) ou des préférences non alignées, comme par exemple dans le cadre de marchés asymétriques ou de conflits déséquilibrés (e.g., marketing entre une startup et un leader établi).

Version pondérée. Chaque champ possède une valeur λ_i spécifique. Par exemple, certaines circonscriptions ou cibles sont plus stratégiques que d'autres dans le domaine politique.

Version à cibles multiples. Au lieu de traiter chaque champ séparément, certains jeux regroupent les cibles en blocs et définissent des règles de victoire à l'échelle du groupe.

1.2.4 Stratégies mixtes

Un peu plus loin, on évoquera les concepts d'équilibres qui seront ce qu'on accepte généralement comme "solution" d'un problème de jeu. Et dans le jeu de Blotto, même s'il est très simple d'énoncé, a certaines variations qui n'admettent pas d'équilibres en s'appuyant uniquement sur les répartitions. Pour passer outre cette limitation, Les joueurs peuvent adopter des stratégies mixtes, c'est-à-dire choisir une distribution de probabilité sur l'ensemble des répartitions admissibles. Une stratégie mixte est donc une mesure de probabilité μ sur l'ensemble Δ_α^m, où :

$$\Delta_\alpha^m = \left\{ x \in \mathbb{R}_+^m \,\middle|\, \sum_{i=1}^{m} x_i = \alpha \right\}$$

Le gain espéré devient :

$$\mathbb{E}_{(X,Y)\sim(\mu,\nu)}\left[G(X,Y)\right]$$

1.3 Résolution du jeu de Blotto

1.3.1 Premiers résultats de Borel

Comme précisé précédemment, Borel n'analyse que quelques cas particuliers. Il note cependant l'intérêt théorique du problème et pose la question de l'existence de stratégies optimales dans des cas plus généraux. Sa formulation introduit déjà des concepts fondamentaux : gains aléatoires, compétition sur plusieurs fronts, et incertitude sur le comportement de l'adversaire.

Borel et Ville [13] parviennent à construire un profil de stratégies d'équilibre dans le cas se restreignant à trois champs de bataille, et à budgets discrets et symétriques. Ils illustrent le potentiel du modèle, mais sans proposer de démonstration d'existence d'équilibres pour des configurations plus générales. Le passage à plus de champs de bataille augmente considérablement la complexité du jeu, et cette question reste, à l'époque, ouverte. Ce problème sera résolu rigoureusement plus tard, notamment par Gross et Wagner.

1.3.2 Modèle continu de Gross et Wagner

Gross et Wagner [31] proposent une extension des solutions du jeu de Blotto dans un cadre continu. Ils donnent un équilibre pour le jeu en version asymétrique à deux champs de bataille, pour le jeu symétrique à trois champs de bataille, et enfin pour le jeu symétrique homogène (c-à-d sans pondération de champs de bataille) pour plusieurs champs de bataille.

Cadre du jeu

Les deux joueurs disposent d'un budget continu de ressources α et β respectivement pour le joueur A et le joueur B. Ces budgets sont à répartir sur m champs de bataille. On note $x \in \mathbb{R}^m$ la stratégie du joueur A, avec :

$$x_i \geq 0, \quad \sum_{i=1}^{3} x_i = 1$$

et $y \in \mathbb{R}^m$ la stratégie du joueur B, satisfaisant la même contrainte.

Le gain local sur chaque champ est donné par :

$$g_i(x_i, y_i) = \begin{cases} 1 & \text{si } x_i > y_i \\ 0.5 & \text{si } x_i = y_i \\ 0 & \text{si } x_i < y_i \end{cases}$$

Le gain global du joueur A est alors :

$$G_A(x, y) = \sum_{i=1}^{m} \lambda_i g_i(x_i, y_i)$$

et le jeu est à *somme nulle* : $G_B = m - G_A$.

Stratégies mixtes et équilibres

Les joueurs peuvent adopter des stratégies mixtes : ils choisissent une loi de probabilité sur l'espace Δ_α^m pour le joueur A et Δ_β^m pour le joueur B, définis en 1.2.4. Leur objectif est alors de maximiser leur gain espéré :

$$\mathbb{E}_{(X,Y)\sim(\mu,\nu)}\left[G_I(X, Y)\right]$$

où μ et ν sont des mesures de probabilité sur Δ_α^m et Δ_β^m respectivement, étant jouées par les joueurs A et B respectivement.

Résultat principal et équilibre de Nash

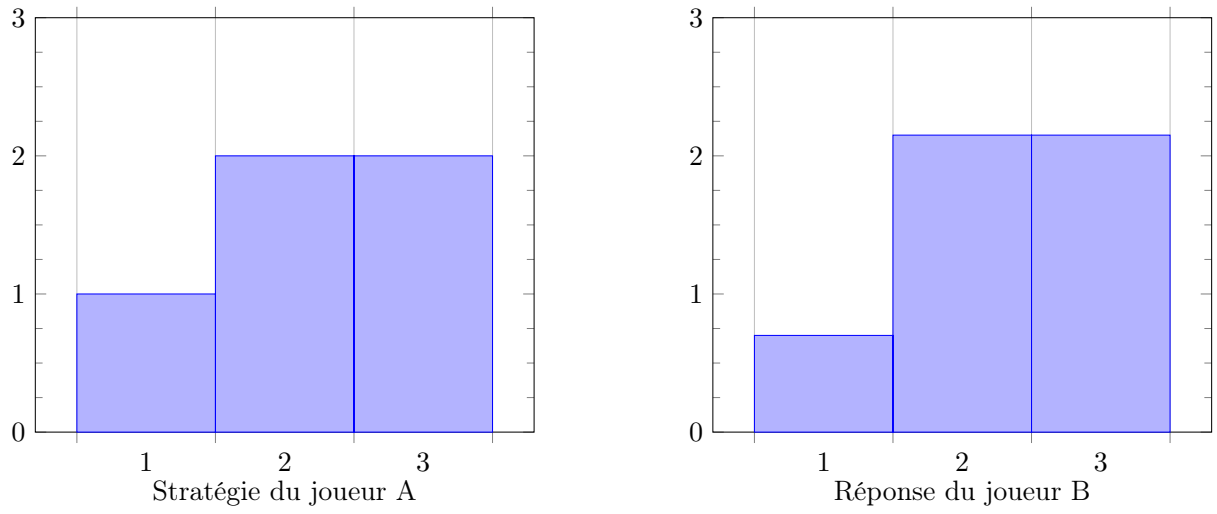
Avant d'énoncer les résultats fondamentaux de Gross et Wagner, on rappelle la notion d'équilibre de Nash dans le cadre du jeu de Blotto.

Définition 1.3.1. *Un profil de stratégies* $(x^*, y^*) \in \Delta_\alpha^m \times \Delta_\beta^m$ *est un* équilibre de Nash en stratégies pures *si aucun joueur des deux joueurs ne peut améliorer son gain en déviant unilatéralement. Autrement dit :*

$$\forall x \in \Delta_\alpha^m, \quad G_A(x, y^*) \leq G_A(x^*, y^*) \quad et \quad \forall y \in \Delta_\beta^m, \quad G_B(x^*, y) \leq G_B(x^*, y^*)$$

Cette notion d'équilibre a été introduite par John F. Nash dans son article fondateur [60], où il démontre l'existence d'un équilibre dans tout jeu fini à n joueurs, en autorisant des distributions de probabilité sur les stratégies pures.

Remarque 1.3.2. *Le jeu de Blotto n'admet pas en général d'équilibre en stratégies pures : on considère l'exemple illustré dans la figure ci-dessous : Un jeu symétrique continu à trois champs de bataille non pondérés. Si le joueur A joue* $x_A = (1, 2, 2)$*, le joueur B peut toujours copier sa stratégie, en enlevant un* ϵ *du champ 1 et en le distribuant sur les champs 2 et 3* $y_B = (1 - \epsilon, 2 + \frac{\epsilon}{2}, 2 + \frac{\epsilon}{2})$*. Dans ce cas, le joueur 2 gagne 2 champs de bataille alors que le joueur 1 n'en gagne qu'un seul. Le joueur 1 peut alors adapter similairement sa stratégie. On aurait alors une boucle infinie de déviations unilatérales lucratives.*



Lorsque l'ensemble des stratégies pures ne permet pas l'existence d'un équilibre, on autorise les joueurs à jouer selon des distributions de probabilité sur l'ensemble des stratégies disponibles. On parle alors d'*équilibre de Nash en stratégies mixtes*.

Dans ce cadre, les stratégies deviennent des mesures de probabilité μ, ν sur les ensembles admissibles, et le gain espéré du joueur A est donné par :

$$\mathbb{E}_{(X,Y)\sim(\mu,\nu)}\left[G_A(X,Y)\right]$$

Définition 1.3.3. *Un couple $(\mu^*, \nu^*) \in \mathcal{P}(\Delta_\alpha^m) \times \mathcal{P}(\Delta_\beta^m)$ constitue un équilibre de Nash en stratégies mixtes si :*

$$\forall \mu, \quad \mathbb{E}_{(X,Y)\sim(\mu,\nu^*)}\left[G(X,Y)\right] \leq \mathbb{E}_{(X,Y)\sim(\mu^*,\nu^*)}\left[G(X,Y)\right]$$

et

$$\forall \nu, \quad \mathbb{E}_{(X,Y)\sim(\mu^*,\nu)}\left[G(X,Y)\right] \geq \mathbb{E}_{(X,Y)\sim(\mu^*,\nu^*)}\left[G(X,Y)\right]$$

où $\mathcal{P}(\Delta_\alpha^m)$ et $\mathcal{P}(\Delta_\beta^m)$ sont respectivement l'ensemble des mesures de probabilité sur Δ_α^m et Δ_β^m.

Théorème 1.3.4 (Gross & Wagner, 1950)**.** *Dans le cadre du jeu de Blotto continu, il existe un équilibre de Nash en stratégies mixtes, caractérisé comme suit :*

- *Si $m = 2$ et $\alpha = \beta$, alors la stratégie consistant à tirer uniformément une proportion $U \sim \mathcal{U}[0,1]$ et à jouer $(\alpha U, \alpha(1-U))$ est optimale pour les deux joueurs. Autrement dit, chaque joueur choisit une allocation uniforme sur le simplexe Δ_α^1.*
- *Si $m = 2$ et $\alpha \neq \beta$, alors l'équilibre est donné par une stratégie mixte discrète, supportée sur un sous-ensemble fini de Δ_α^1. Dans ce cas, la valeur du jeu peut être exprimée explicitement en fonction de α, β et des valeurs des champs de bataille.*
- *Si $m = 3$ et $\alpha = \beta$, la distribution uniforme sur Δ_α^2 est optimale pour les deux joueurs.*
- *Plus généralement, si $m \geq 3$, $\alpha = \beta$ et $\lambda_i = \lambda_j$ pour tout $i, j = 1, \ldots, m$, alors la distribution uniforme sur Δ_α^{m-1} constitue une stratégie optimale pour les deux joueurs.*

Structure des stratégies et portée du résultat

Ce résultat met en évidence une dichotomie entre les cas $m = 2$ et $m \geq 3$. Lorsque $m = 2$ et que les budgets sont égaux, l'équilibre consiste à tirer uniformément une proportion de budget et à l'allouer au premier champ, le second recevant le complément. Géométriquement, cela revient à une distribution uniforme sur un simplexe de dimension 1, c'est-à-dire un intervalle. Lorsque $m = 2$ mais que les budgets diffèrent, la symétrie est rompue : l'équilibre n'est plus donné par une loi continue mais par une stratégie mixte discrète, supportée sur un ensemble fini de points du simplexe. En revanche, dès que $m \geq 3$ et que tous les champs sont de même valeur, l'équilibre retrouve une structure continue : il s'agit de la distribution uniforme sur le simplexe Δ_α^{m-1}, qui généralise le cas $m = 2$ à des dimensions supérieures.

Gross et Wagner n'arrivent cependant pas à trouver des expressions fermées dans tous les cas de figure. Leur construction repose sur la symétrie parfaite des champs de bataille et sur la géométrie particulière du cas à trois champs. Pour des valeurs plus générales de m, ou en présence de pondérations λ_i, la structure des équilibres devient nettement plus complexe.

Ce travail a servi de point de départ à une riche littérature sur les variantes du jeu de Blotto. On reviendra dans la section suivante sur certaines de ces extensions, en particulier dans des cadres asymétriques, pondérés ou dynamiques.

1.3.3 Avancées sur le jeu de Blotto

Depuis les travaux de Gross et Wagner, la variante continue du jeu du Colonel Blotto est aujourd'hui bien mieux comprise. En utilisant une approche fondée sur la théorie des copules, Roberson [68] a achevé la solution des cas asymétriques. Plus précisément, il a obtenu une caractérisation complète des distributions marginales des stratégies d'équilibre. Macdonell et Mastronardi [58] ont obtenu une caractérisation complète des équilibres de Nash dans le cas de deux champs de bataille. Ces avancées ont posé les bases de l'étude de diverses extensions du jeu de Colonel Blotto, telles que la variante non à somme nulle étudiée par Roberson et Kvasov [69] ; la généralisation permettant des valeurs hétérogènes des champs de bataille, y compris la variante non à somme nulle où les valeurs sont hétérogènes entre les joueurs, étudiée par Kovenock et Roberson [52] ; l'extension à plus de deux joueurs, étudiée par Boix-Adserà et al. [11] ; et l'extension autorisant des *head-starts* étudiée par Vu et Loiseau [80].

Pour la variante discrète, la contribution la plus importante pour obtenir une solution explicite du jeu est le travail de Hart [38]. En introduisant une variante du jeu, appelée jeu du *General Lotto*, qui relaxe la condition de budget en condition de budget espéré, Hart a trouvé des stratégies optimales ainsi que la valeur du jeu de Colonel Blotto discret dans le cas symétrique, où les deux joueurs ont le même nombre d'unités, ainsi que dans les cas asymétriques où les budgets des joueurs sont proches. Il a également obtenu des bornes inférieures ou supérieures sur la valeur du jeu (gain du joueur A) dans certains cas.

Ensuite, Dziubiński [24] a construit des stratégies optimales supplémentaires pour le cas symétrique et les a utilisées pour obtenir une caractérisation complète d'une statistique agrégée des stratégies optimales dans ce cas, appelée spectre. Dans un article plus récent, Liang et al. [56] ont étudié le jeu de Colonel Blotto discret avec deux champs de bataille. Ils ont obtenu une caractérisation complète des valeurs du jeu, fourni une caractérisation directe complète des stratégies optimales des joueurs dans les cas où la différence de ressources entre les joueurs est $\beta = \alpha - 1$, ainsi que dans les cas où $\beta \leq \alpha - 2$ et $\alpha \bmod (\alpha - \beta) = \alpha - \beta - 1$. Dans les cas restants, ils ont fourni une caractérisation complète mais essentiellement indirecte, où les stratégies de l'un des joueurs sont décrites en termes d'un ensemble de conditions. Et, très récemment, Dziubiński [25] résout le cas à trois champs de batailles ou plus, avec la contrainte de budget $\beta = (\alpha \bmod m)$. C'est d'ailleurs de ce dernier article qu'on s'inspire pour établir cet inventaire de résultats.

D'un point de vue algorithmique, le problème du calcul de stratégies optimales dans le jeu de Colonel Blotto a attiré l'attention de chercheurs en recherche opérationnelle et en informatique pendant de nombreuses années. Beale et Heselden [5] ont proposé une méthode pour trouver des solutions approximatives du jeu de Colonel Blotto, fondée sur un jeu auxiliaire dans lequel les contraintes de budget des joueurs doivent être satisfaites seulement en espérance. En utilisant partiellement cette approche, Penn [66] a proposé une méthode de multiplicateur généralisé qui peut être appliquée à une large classe de jeux, incluant le jeu de Colonel Blotto. Cependant, aucun de ces articles ne s'intéresse à la complexité calculatoire du problème.

Une percée sur ce point a été introduite par l'article d'Ahmadinejad et al. [1], qui ont montré que le jeu de Colonel Blotto (et des types de jeux similaires avec gains bilinéaires) peut être résolu en temps polynomial par rapport aux valeurs des paramètres du jeu. Leur méthode, bien que polynomiale, a un coût de calcul élevé car elle nécessite d'utiliser la méthode de l'ellipsoïde pour résoudre le programme linéaire associé à la solution du jeu. Cette approche a été signifi-

cativement améliorée par Behnezhad et al. [6] qui, en utilisant une représentation des stratégies mixtes basée sur les flots dans un réseau pour les jeux de Colonel Blotto, ont construit un programme linéaire pour résoudre le jeu, suffisamment petit pour permettre un calcul pratique.

Enfin, Boix-Adserà et al. [11] ont proposé un algorithme d'échantillonnage qui permet de générer des stratégies mixtes dans des jeux de Colonel Blotto multijoueurs à partir de distributions marginales données. Vu et Loiseau [80] ont proposé un algorithme efficace pour calculer de bonnes (mais pas nécessairement d'équilibre) distributions marginales dans le jeu de Colonel Blotto avec favoritisme. Ces distributions marginales peuvent ensuite être utilisées pour obtenir des équilibres de Nash approximatifs dans le jeu. Beaglehole et al. [4] ont proposé une extension rapide par échantillonnage de l'algorithme de mise à jour par poids (*Multiplicative Weights Update*) et ont montré comment celui-ci peut être utilisé pour résoudre rapidement le jeu de Colonel Blotto.

1.4 Jeux différentiels stochastiques et extension du jeu de Blotto

On passe à présent aux jeux dans le cadre différentiel stochastique, c'est-à-dire des jeux qui évoluent en temps continu et qui impliquent du hasard dans leur évolution ou leur conclusion. L'objectif de cette section est d'introduire l'analyse mathématique des jeux différentiels stochastiques. On s'appuiera fortement dans cette section sur l'ouvrage de Carmona [16], auquel on pourra se référer pour une introduction plus complète au contrôle stochastique et aux jeux différentiels stochastiques.

1.4.1 Formulation des jeux différentiels stochastiques

On considère N joueurs, notés $1, \dots, N$, et on suppose qu'à chaque instant t, les joueurs agissent sur un système dont l'état X_t évolue sous l'influence de leurs décisions, jusqu'à un temps terminal T. La dynamique de ce système est gouvernée par une équation différentielle stochastique au sens d'Itô, qu'on donne dans l'équation 1.1. Plus précisément, la dynamique est pilotée par un mouvement brownien m-dimensionnel $(W_t)_{0\le t\le T}$ défini sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, avec filtration naturelle $(\mathcal{F}_t)_{0\le t\le T}$.

La différence principale avec le contrôle stochastique réside dans les fonctions de gains, dans le sens où chaque joueur optimise un critère qui pourrait être antagoniste aux autres. Chaque joueur $i \in \{1, \dots, N\}$ dispose d'un ensemble d'actions A^i, en général un espace métrique compact ou un sous-ensemble de $\mathbb{R}^{k_i}$ muni de sa tribu borélienne. On note $\mathcal{A}^i$ l'ensemble des processus adaptés à valeurs dans A^i, représentant les stratégies admissibles du joueur i. L'ensemble des profils de stratégies est alors le produit $\mathcal{A} = \mathcal{A}^1 \times \cdots \times \mathcal{A}^N$.

Un profil de stratégies $a = (a^1, \dots, a^N) \in \mathcal{A}$ est constitué de N processus stochastiques adaptés $(a^i_t)_{0\le t\le T}$, qui doivent satisfaire certaines conditions supplémentaires, comme des contraintes d'intégrabilité ou encore des restrictions spécifiques à chaque application. De manière générale, ces conditions garantissent que les stratégies des différents joueurs sont bien définies et que le jeu est bien posé.

Pour chaque profil de stratégies $a = (a_t)_{0\le t\le T} \in \mathcal{A}$, on suppose que l'évolution temporelle de l'état $X = X^a$ du système est donnée par l'équation différentielle stochastique suivante :

$$\begin{cases} dX_t = b(t, X_t, a_t)\, dt + \sigma(t, X_t, a_t)\, dW_t, \quad 0 \le t \le T, \\ X_0 = x, \end{cases} \tag{1.1}$$

où les fonctions

$$b : [0,T] \times \Omega \times \mathbb{R}^m \times \mathcal{A} \longrightarrow \mathbb{R}^m, \qquad \sigma : [0,T] \times \Omega \times \mathbb{R}^m \times \mathcal{A} \longrightarrow \mathbb{R}^{m\times m}$$

satisfont les conditions suivantes :

(H1) Pour tout $x \in \mathbb{R}^m$ et tout $a \in \mathcal{A}$, les processus $(b(t,x,a))_{0\le t\le T}$ et $(\sigma(t,x,a))_{0\le t\le T}$ sont progressivement mesurables, à valeurs dans $\mathbb{R}^m$ et $\mathbb{R}^{m\times m}$ respectivement.

(H2) Il existe une constante $c > 0$ telle que, pour tout $t \in [0,T]$, pour tout $a \in \mathcal{A}$, pour tout $\omega \in \Omega$, et pour tous $x, x' \in \mathbb{R}^m$, on ait :

$$|b(t,\omega,x,a) - b(t,\omega,x',a)| + \|\sigma(t,\omega,x,a) - \sigma(t,\omega,x',a)\| \le c\,|x - x'|.$$

Ces hypothèses garantissent l'existence (et l'unicité) d'une solution forte de (1.1).

1.4.2 Les gains des joueurs

On suppose que chaque joueur reçoit des gains instantanés et terminaux. Chaque joueur décidera alors de sa stratégie afin d'optimiser son gain total. Plus précisément, pour chaque joueur $i \in \{1, \dots, N\}$, on considère :

— une variable aléatoire $\mathcal{F}_T$-mesurable $\xi^i \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, appelée *gain terminal*. Dans de nombreux cas, ce gain s'écrit sous la forme $\xi^i = g^i(X_T)$, où $g^i : \Omega \times \mathbb{R}^d \to \mathbb{R}$ est une fonction mesurable, supposée Lipschitz ou à croissance au plus quadratique ;

— une fonction $f^i : [0,T] \times \Omega \times \mathbb{R}^d \times \mathcal{A} \to \mathbb{R}$, appelée *gain instantané*, satisfaisant les mêmes hypothèses de régularité que le drift b. Dans de nombreux modèles appliqués, ce gain est pondéré par un facteur d'actualisation exponentiel $e^{-\rho t}$, avec $\rho > 0$, de manière à refléter le fait que les gains ou pertes immédiats pèsent davantage que ceux réalisés dans le futur.

Le gain total attendu pour le joueur i, lorsque les N joueurs suivent un profil de stratégies $a \in \mathcal{A}$, est alors défini par la fonctionnelle de gain :

$$J^i(a) = \mathbb{E}\left[\int_0^T f^i(s, X_s, a_s)\, ds + \xi^i\right], \qquad a = (a^1, \dots, a^N) \in \mathcal{A}, \tag{1.2}$$

où X désigne le processus d'état du système, solution de l'équation (1.1).

En général, le gain d'un joueur dépend indirectement des stratégies choisies par les autres via l'évolution de l'état X_t, mais aussi directement puisque les actions a_t^j des autres joueurs peuvent apparaître explicitement dans l'expression du gain instantané f^i.

Chaque joueur cherche alors à maximiser son gain total J^i. Si l'on regroupe l'ensemble des gains sous la forme

$$J(a) = \big(J^1(a), \dots, J^N(a)\big), \qquad a \in \mathcal{A}$$

on obtient la fonctionnelle de gain associée au jeu différentiel stochastique à N joueurs.

Définition 1.4.1 (Équilibre de Nash)**.** *Un profil de stratégies admissibles*

$$a^\star = (a^{1,\star}, \dots, a^{N,\star}) \in \mathcal{A} = \mathcal{A}^1 \times \cdots \times \mathcal{A}^N$$

est un équilibre de Nash *pour le jeu si, pour tout $i \in \{1, \dots, N\}$ et pour toute stratégie admissible $a^i \in \mathcal{A}^i$,*

$$J^i\big(a^{1,\star}, \dots, a^{i-1,\star}, a^{i,\star}, a^{i+1,\star}, \dots, a^{N,\star}\big) \;\geq\; J^i\big(a^{1,\star}, \dots, a^{i-1,\star}, a^i, a^{i+1,\star}, \dots, a^{N,\star}\big).$$

Autrement dit, aucune déviation unilatérale de $a^{i,\star}$ n'améliore le gain du joueur i.

Remarque 1.4.2. *Un équilibre de Nash n'est en général pas optimal au sens de Pareto, c'est-à-dire qu'il peut exister d'autres profils de stratégies qui améliorent la situation d'au moins un joueur sans détériorer celle des autres. Cette amélioration nécessite cependant une déviation de plus d'un joueur.*

Remarque 1.4.3. *Un équilibre de Nash n'est pas unique en général, et en particulier dans le cas des jeux différentiels stochastiques.*

Remarque 1.4.4. *Si on regarde le jeu du point de vue d'un joueur i à profil fixe de stratégies des autres joueurs, le problème est en fait un problème de contrôle stochastique. C'est cette perspective-là qu'on adoptera pour déterminer les équilibres de Nash dans la partie suivante.*

1.4.3 Jeux différentiels stochastiques markoviens

Dans un grand nombre d'applications, on se retrouve avec des jeux markoviens, c'est-à-dire que les fonctions de dérive b et de diffusion σ sont markoviennes : elles ne dépendent à un instant t donné de X et W que par leurs valeurs respectives au même instant X_t et W_t. Cette structure est d'autant plus intéressante car elle permet de faire des résolutions numériques.

On supposera de plus que les fonctions de gain instantané f^i et les variables aléatoires de gain terminal ξ^i sont markoviennes, au sens où, comme b et σ, la fonction f^i ne dépend pas du scénario aléatoire $\omega \in \Omega$, mais uniquement de la valeur courante de l'état X et des actions entreprises par les joueurs. Ainsi, on peut écrire

$$f^i : [0,T] \times \mathbb{R}^d \times \mathcal{A} \;\to\; \mathbb{R}, \qquad (t,x,a) \mapsto f^i(t,x,a) \in \mathbb{R},$$

et le gain terminal sous la forme $\xi^i = g^i(X_T)$ pour une certaine fonction mesurable $g^i : \mathbb{R}^m \to \mathbb{R}$, généralement Lipschitz ou à croissance (au plus) quadratique.

Dans ce cadre markovien, la fonctionnelle de gain du joueur i prend la forme

$$J^i(a) = \mathbb{E}\left[\int_0^T f^i(t, X_t, a_t)\, dt + g^i(X_T)\right], \qquad a \in \mathcal{A}. \tag{1.3}$$

Cette thèse couvre des jeux markoviens et non markoviens ; dans le second cas, des outils comme les équations différentielles stochastiques rétrogrades permettent de considérer des critères dépendant de toute la trajectoire du processus d'état. Ces notions sont précisées plus loin dans le manuscrit.

Remarque 1.4.5. *Dans le cadre markovien, certaines stratégies optimales constituant des équilibres de Nash peuvent être représentées comme des fonctions déterministes de l'état du jeu. On les désigne alors sous le nom d'*équilibres markoviens de Nash*, ou d'*équilibres de Nash en stratégies feedback*. Par abus de langage, on utilisera parfois le terme d'« équilibre » ou de « stratégies optimales » pour faire référence soit directement au profil de stratégies soit aux fonctions qui caractérisent ces dernières.*

1.4.4 Approche par les EDP

Dans un cadre markovien, il est naturel de recourir à l'approche par équations aux dérivées partielles pour traiter les problèmes de contrôle optimal intervenant dans la construction des stratégies optimales. Cette approche s'appuie sur les équations de Hamilton–Jacobi–Bellman (HJB). Pour ce faire, on introduit, pour toute stratégie a, l'opérateur suivant :

$$\mathcal{L}_t^a = \tfrac{1}{2} \sum_{h,k=1}^{m} \ell_{hk}^a(t,x)\, \frac{\partial^2}{\partial x_h \partial x_k} + \sum_{h=1}^{m} b_h^a(t,x)\, \frac{\partial}{\partial x_h}\,.$$

Cet opérateur est le générateur différentiel du second ordre associé au processus d'état sous la stratégie a. Il traduit l'action de la dynamique contrôlée sur les fonctions suffisamment régulières de l'état (t,x).

Les coefficients b et l sont définis comme suit :

$$\begin{cases} b_h^a(t,x) \;=\; b_h\big(t,x,a\big), & h=1,\dots,m,\\ \ell_{hk}^a(t,x) \;=\; \displaystyle\sum_{j=1}^{m} \sigma_{hj}\big(t,x,a\big)\,\sigma_{kj}\big(t,x,a\big), & h,k=1,\dots,m. \end{cases}$$

Définition 1.4.6. *Pour chaque joueur $i \in \{1,\dots,N\}$ et pour un profil de stratégies a, on définit la fonction de valeur V^i par*

$$V^i(t,x) = \sup_{a^i\in\mathcal{A}^i} \mathbb{E}\bigg[\int_t^T f^i\big(s,X_s,(a^{-i}(t,x),a^i)\big)\,ds + g^i(X_T)\;\Big|\; X_t=x\bigg].$$

où a^{-i} correspond à l'ensemble des stratégies jouées par les joueurs différents de i.

Cette fonction correspond au gain du joueur i s'il joue de manière optimale face à des stratégies figées a^{-i}.

Le lien avec les EDP se fait grâce à cette fonction, qui, dans le cas d'un seul joueur, traduit un problème de contrôle stochastique.

Théorème 1.4.7 (Réponse optimale et équation HJB)**.** *Face à un profil de stratégies (a^{-i}), la fonction valeur V^i traduisant le gain associé à la réponse optimale du joueur i est une solution de viscosité de l'EDP d'Hamilton-Jacobi-Bellman suivante :*

$$\begin{cases} \partial_t\phi^i(t,x) + \sup\limits_{a^i\in A^i}\Big\{f^i\big(t,x,(a^{-i}(t,x),a^i)\big) + \mathcal{L}_t^{(a^{-i}(t,x),a^i)}\,\phi^i(t,x)\Big\} = 0, \quad t<T,\\ \phi^i(T,x) = g^i(x). \end{cases} \tag{1.4}$$

La stratégie optimale est alors donnée à chaque état du jeu (t,x) par :

$$a^{i,\star}(t,x) \in \arg\max_{a^i\in A^i}\Big\{f^i\big(t,x,a^{-i}(t,x),\,a^i\big) + \mathcal{L}_t^{\left(a^{-i}(t,x),\,a^i\right)}\,V^i(t,x)\Big\}. \tag{1.5}$$

On trouve une démonstration de ce résultat dans la section contrôle stochastique de l'ouvrage de Carmona [16]

Remarque 1.4.8. *Dans le théorème 1.4.7, on a a priori envie d'utiliser une formule d'Itô sur V^i, sauf qu'on n'a aucune garantie de régularité sur cette dernière. C'est la raison pour laquelle on passe par des fonctions test et on appelle v solution de viscosité. On renvoie à Crandall et. al. [19] pour plus de détails.*

Remarque 1.4.9. *Il convient de distinguer soigneusement deux niveaux d'optimisation. Dans la formulation probabiliste du jeu stochastique, chaque joueur choisit un processus de contrôle $(a_t^i)_{0\le t\le T}$ adapté à la filtration, et l'optimisation porte sur l'ensemble de ces processus admissibles. En revanche, dans la formulation en équations aux dérivées partielles (HJB), le* principe de programmation dynamique *permet de ramener l'optimisation à un problème* ponctuel *en chaque paire (t,x) : l'équation ne dépend que de la valeur instantanée $a^i \in A^i$ que peut prendre la stratégie au point (t,x). Ainsi, l'optimisation dans l'EDP se simplifie : elle se fait sur l'espace des* valeurs de stratégies *A^i à dimension finie, et non plus sur l'ensemble des trajectoires complètes de contrôles $\mathcal{A}^i$ à dimension infinie.*

Ainsi, chaque joueur i associe à son problème de réponse optimale une fonction valeur V^i qui satisfait une équation de type Hamilton–Jacobi–Bellman. Si l'on parvient à identifier une famille de fonctions $(V^1, \dots, V^N)$ et de stratégies en feedback $(a^{1,\star}, \dots, a^{N,\star})$ telles que, pour chaque joueur i, la stratégie $a^{i,\star}$ réalise effectivement l'optimisation dans son équation HJB, alors aucun joueur n'a intérêt à dévier unilatéralement. On obtient dans ce cas un *équilibre de Nash markovien*, c'est-à-dire une configuration stable où chaque joueur joue une stratégie optimale vis-à-vis de celles des autres.

Théorème 1.4.10 (Équilibre de Nash markovien via HJB)**.** *On suppose qu'il existe des fonctions de valeur $V^i : [0,T] \times \mathbb{R}^m \to \mathbb{R}$ et des stratégies $a^{i,\star} : [0,T] \times \mathbb{R}^m \to A^i$ mesurables telles que, pour tout $i \in \{1, \dots, N\}$, V^i est solution de viscosité de l'EDP :*

$$\partial_t V^i(t,x) \; + \; \max_{a^i \in A^i} \Big\{ f^i\big(t,x,((a^\star)^{-i}(t,x),a^i)\big) + \mathcal{L}_t^{((a^\star)^{-i}(t,x),a^i)} V^i(t,x) \Big\} = 0, \quad t < T, \tag{1.6}$$

avec condition terminale

$$V^i(T,x) = g^i(x), \tag{1.7}$$

et où le maximum de (1.6) *est atteint au point $a^{i,\star}(t,x)$, c'est-à-dire*

$$a^{i,\star}(t,x) \in \arg\max_{a^i \in A^i} \Big\{ f^i\big(t,x,((a^\star)^{-i}(t,x),a^i)\big) + \mathcal{L}_t^{((a^\star)^{-i}(t,x),a^i)} V^i(t,x) \Big\}. \tag{1.8}$$

Alors le profil de stratégies $a^\star = (a^{1,\star}, \dots, a^{N,\star})$ est un équilibre de Nash markovien : pour tout i et toute stratégie admissible $\tilde{a}^i \in \mathcal{A}^i$,

$$J^i\big((a^\star)^{-i}, a^{i,\star}\big) \;\geq\; J^i\big((a^\star)^{-i}, \tilde{a}^i\big),$$

et V^i coïncide avec la valeur maximale du gain associée à $((a^\star)^{-i}, a^{i,\star})$.

Ce résultat, dont une démonstration détaillée est donnée par Touzi [77], établit le lien entre jeux différentiels stochastiques markoviens et équations de Hamilton-Jacobi-Bellman.
On s'intéresse dans la section suivante à un cas particulier de ces jeux, les jeux à somme nulle, auquel se rattache l'une des contributions de cette thèse sous la forme d'une extension du jeu de Blotto.

1.4.5 Jeux à somme nulle

Les jeux différentiels stochastiques à deux joueurs et à somme nulle suscitent un intérêt particulier dans la mesure où ils modélisent de nombreuses situations réelles et se prêtent plutôt bien à la résolution numérique. Ils permettent parfois de trouver des solutions explicites pour les fonctions valeurs (et donc les espérances de gain), ainsi que pour les stratégies optimales menant à un équilibre de Nash.

Dans le cadre de ce chapitre, un jeu à somme nulle est un jeu à deux joueurs $N = 2$ pour lequel $J^1 + J^2 = 0$, auquel cas on se concentre uniquement sur la fonction objectif $J = J^1 = -J^2$ du premier joueur. Autrement dit, le joueur 1 cherche à maximiser le gain $J(a^1, a^2)$, tandis que le joueur 2 cherche à minimiser cette même quantité.

Proposition 1.4.11. *Pour un jeu à somme nulle, un profil de stratégies $a \in \mathcal{A}$ est un point selle si*

$$\sup_{a^1 \in \mathcal{A}^1} \inf_{a^2 \in \mathcal{A}^2} J(a^1, a^2) = \inf_{a^2 \in \mathcal{A}^2} \sup_{a^1 \in \mathcal{A}^1} J(a^1, a^2). \tag{1.9}$$

Remarque 1.4.12. *Tout point selle est un équilibre de Nash. En effet, si une action constitue un optimum pour un joueur indépendamment de l'action choisie par l'adversaire, alors elle est en particulier une meilleure réponse à toute stratégie de ce dernier, et donc à une action optimale de l'autre joueur.*

Définition 1.4.13. *L'étude des jeux à somme nulle repose sur l'analyse des valeurs inférieure et supérieure du jeu, définies respectivement par*

$$\underline{V}_0 = \sup_{a^1 \in \mathcal{A}^1} \inf_{a^2 \in \mathcal{A}^2} J(a^1, a^2), \qquad \overline{V}_0 = \inf_{a^2 \in \mathcal{A}^2} \sup_{a^1 \in \mathcal{A}^1} J(a^1, a^2). \tag{1.10}$$

On appelle valeur du jeu, et l'on note V_0, la valeur commune lorsque les deux quantités coïncident, c'est-à-dire lorsque

$$\underline{V}_0 = \overline{V}_0 = V_0.$$

On vérifie immédiatement que $\underline{V}_0 \leq \overline{V}_0$, ce qui justifie leur terminologie. L'analyse d'un jeu à somme nulle se concentre généralement sur les deux questions suivantes :

(i) Quand la valeur du jeu existe-t-elle ?

(ii) Lorsque la valeur du jeu existe, peut-on trouver des stratégies optimales formant un point selle ? En d'autres termes, existe-t-il $a^{1,\star} \in \mathcal{A}^1$ et $a^{2,\star} \in \mathcal{A}^2$ tels que

$$V_0 = J(a^{1,\star}, a^{2,\star}) = \sup_{a^1 \in \mathcal{A}^1} J(a^1, a^{2,\star}) = \inf_{a^2 \in \mathcal{A}^2} J(a^{1,\star}, a^2). \tag{1.11}$$

Définition 1.4.14 (Condition d'Isaacs)**.** *On dit que la condition d'Isaacs est satisfaite si, pour tout $(t,x) \in [0,T] \times \mathbb{R}^m$ et toute fonction test $\varphi \in C^2(\mathbb{R}^m)$, on a*

$$\sup_{a^1 \in A^1} \inf_{a^2 \in A^2} \left\{ f^1(t,x,a^1,a^2) + \mathcal{L}_t^{(a^1,a^2)} \varphi(x) \right\} = \inf_{a^2 \in A^2} \sup_{a^1 \in A^1} \left\{ f^1(t,x,a^1,a^2) + \mathcal{L}_t^{(a^1,a^2)} \varphi(x) \right\}. \tag{1.12}$$

Le lien entre les valeurs d'un jeu stochastique à somme nulle et les équations aux dérivées partielles se traduit par l'introduction des équations de Hamilton–Jacobi–Bellman–Isaacs (HJBI), qui permettent de caractériser les fonctions valeurs comme solutions de viscosité d'EDPs adaptées.

Définition 1.4.15. *On appelle équations de Hamilton–Jacobi–Bellman–Isaacs (HJBI) les EDPs*

$$\partial_t \overline{V}(t,x) + \inf_{a^2 \in A^2} \sup_{a^1 \in A^1} \left\{ f^1(t,x,a^1,a^2) + \mathcal{L}_t^{(a^1,a^2)} \overline{V}(t,x) \right\} = 0, \tag{1.13}$$

et

$$\partial_t \underline{V}(t,x) + \sup_{a^1 \in A^1} \inf_{a^2 \in A^2} \left\{ f^1(t,x,a^1,a^2) + \mathcal{L}_t^{(a^1,a^2)} \underline{V}(t,x) \right\} = 0, \tag{1.14}$$

avec condition terminale

$$\overline{V}(T,x) = g^1(x).$$

L'équation (1.13) *est dite* équation de valeur supérieure *et* (1.14) *est dite* équation de valeur inférieure.

Proposition 1.4.16. *La valeur supérieure $\overline{V}$ d'un jeu stochastique à somme nulle est une sous-solution de viscosité de l'équation* (1.13)*, tandis que la valeur inférieure $\underline{V}$ est une sur-solution de viscosité de l'équation* (1.14)*.*

Lorsque la condition d'Isaacs (1.12) *est satisfaite, les deux équations coïncident et l'on a*

$$V = \underline{V} = \overline{V},$$

qui est alors une solution de viscosité de l'EDP de Hamilton–Jacobi–Bellman–Isaacs commune.

Une démonstration rigoureuse de ce résultat peut être trouvée dans Touzi [77].

Remarque 1.4.17. *L'existence d'une solution commune aux deux équations HJBI équivaut à l'existence d'une valeur du jeu sous forme de solution de viscosité de l'EDP HJBI. Cela garantit en particulier que les joueurs disposent de stratégies optimales menant à un équilibre de Nash markovien.*

1.4.6 Contribution : jeu de Blotto différentiel stochastique symétrique

On propose dans cette thèse une extension du jeu de Blotto standard au cadre différentiel stochastique, la formulation se prête naturellement au cadre de somme nulle. Dans ce travail, on propose plusieurs extensions du jeu classique de Colonel Blotto dans différentes directions.

— On introduit tout d'abord une version dynamique en temps continu du jeu de Blotto, dans laquelle les deux joueurs peuvent observer l'évolution des processus d'état et ajuster en continu leurs allocations. On définit en particulier une version stochastique de ce jeu, où les joueurs contrôlent la dérive de la solution d'une équation différentielle stochastique (EDS).

— On considère ensuite des critères généraux : les hypothèses formulées sur la fonction de gain associée à chaque champ de bataille sont relativement faibles.

— Enfin, on autorise des interactions entre champs de bataille. Dans la formulation standard du jeu de Blotto, le gain global d'un joueur est donné par la somme des gains sur chaque champ, ce qui implique que le gain à l'équilibre ne dépend que des distributions marginales des allocations jouées par les joueurs. Dans notre cadre, on considère des critères plus généraux, ouvrant ainsi la voie à un champ d'applications plus large.

Formulation et résolution du jeu

On considère un horizon fini $T > 0$ et $m \in \mathbb{N}^*$ champs de bataille. Le système évolue sur l'espace canonique des trajectoires continues à valeurs dans $\mathbb{R}^m \times \mathbb{R}^m$, muni de sa tribu borélienne et de la filtration canonique. Les deux joueurs disposent de budgets $\alpha, \beta > 0$, et choisissent des processus progressifs $(a_t)_{0\le t\le T}$ et $(b_t)_{0\le t\le T}$ représentant leurs allocations admissibles.

Le jeu est formulé de manière faible, avec un changement de mesure de probabilité associé aux contrôles (a, b). Sous la probabilité induite, les processus dynamiques des allocations des deux joueurs s'écrivent comme des équations différentielles stochastiques pilotées par des mouvements browniens :

$$\begin{aligned} X_s^{1,t,x_1} &= x_1 + \int_t^s a_u du + \int_t^s dW_u^{1,a}, \\ X_s^{2,t,x_2} &= x_2 + \int_t^s b_u du + \int_t^s dW_u^{2,b}. \end{aligned} \tag{1.15}$$

Pour plus de détails sur la construction, on renvoie au chapitre 2.

On considère une fonction de gain $g : \mathbb{R}^m \to \mathbb{R}$, qui dépend uniquement de la différence entre les valeurs finales des processus des deux joueurs.

On supposera que g satisfait les hypothèses suivantes :

(i) g admet une croissance au plus exponentielle, c'est-à-dire qu'il existe des constantes C_1 et C_2, strictement positives, telles que :

$$|g(x)| \leq C_1 e^{C_2|x|}, \quad \forall x \in \mathbb{R}^m$$

(ii) l'ensemble des points de discontinuité de g est de mesure de Lebesgue nulle.

Le critère du jeu est alors défini, pour un état $(x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^m$ à l'instant t et des contrôles (a, b), par

$$J(t, x_1, x_2, a, b) \;=\; \mathbb{E}^{P^{a,b}}\Big[\, g\Big(X_T^{1,t,x_1} - X_T^{2,t,x_2}\Big)\Big]$$

où, sous la mesure $P^{a,b}$, les dynamiques $(X^{1,t,x_1}, X^{2,t,x_2})$ sont données par les équations 1.15.

On considère désormais le jeu où les deux joueurs disposent du même budget. Sans perte de généralité, on suppose $\alpha = \beta = 1$ pour la suite. On a donc en particulier $\mathcal{A} = \mathcal{B}$ et les contrôles sont pris dans $\mathcal{A} := \mathcal{A}^1$.

Un équilibre de Nash dans ce jeu peut être obtenu en résolvant deux problèmes d'optimisation, selon qui joue en premier. Un scénario correspond à Joueur 1 maximisant le critère en fonction de la meilleure réponse de Joueur 2, tandis que l'autre correspond à Joueur 2 minimisant le critère en fonction de la meilleure action de Joueur 1. Pour un état de jeu (x_1, x_2) et un instant t, on obtient :

$$\sup_{a\in\mathcal{A}} \inf_{b\in\mathcal{A}} J(t, x_1, x_2, a, b) \quad \text{(Joueur 1 joue en premier)}, \tag{1.16}$$

$$\inf_{b\in\mathcal{A}} \sup_{a\in\mathcal{A}} J(t, x_1, x_2, a, b) \quad \text{(Joueur 2 joue en premier)}. \tag{1.17}$$

Afin de résoudre ces problèmes d'optimisation, on utilise l'approche de la programmation dynamique, qui conduit à l'analyse des équations d'Isaacs supérieures et inférieures.

Définition 1.4.18. *On définit la fonction w sur $[0, T] \times \mathbb{R}^m$ par*

$$w(t, x) := \frac{1}{(4\pi(T-t))^{\frac{m}{2}}} \int_{\mathbb{R}^m} \exp\left[-\frac{|x-z|^2}{4(T-t)}\right] g(z)dz \tag{1.18}$$

La fonction w est de classe $\mathcal{C}^{1,2}([0, T] \times \mathbb{R}^m)$ et elle est solution de l'équation de la chaleur rétrograde avec condition terminale donnée par g :

$$\partial_t w + \Delta w = 0 \quad \textit{sur } [0, T) \times \mathbb{R}^m, \qquad w(T, x) = g(x) \quad \textit{sur } \mathbb{R}^m.$$

Remarque 1.4.19. *La fonction w peut être interprétée comme la régularisation de la fonction gain terminale g par le noyau de la chaleur. Autrement dit, $w(t, \cdot)$ représente la diffusion rétrograde de g dans le temps, ce qui traduit le fait que les gains espérés des joueurs évoluent en fonction de la dynamique aléatoire du système.*

Définition 1.4.20. *Soit $(t, x_1, x_2) \in [0,T) \times \mathbb{R}^m \times \mathbb{R}^m$. On définit*

$$\tilde{a}(t,x_1,x_2) := \begin{cases} e_i \text{ si } \max\limits_{j=1,\dots,m} \partial_j w = \partial_i w \text{ et } \max\limits_{j=1,\dots,m} \partial_j w > 0 \\ 0_{\mathbb{R}^m} \text{ sinon} \end{cases}$$

et

$$\tilde{b}(t,x_1,x_2) := \begin{cases} e_i \text{ si } \min\limits_{j=1,\dots,m} -\partial_j w = -\partial_i w \text{ et } \min\limits_{j=1,\dots,m} -\partial_j w < 0 \\ 0_{\mathbb{R}^m} \text{ sinon} \end{cases}$$

où, pour tout $i = 1, \dots, m$, $e_i \in \mathbb{R}^m$ désigne le i-ème vecteur de la base canonique de $\mathbb{R}^m$, et $\partial_i w$ est la dérivée partielle de w par rapport à la i-ème coordonnée de l'espace, évaluée en $(t, x_1 - x_2) \in [0,T] \times \mathbb{R}^m$.

Définition 1.4.21. *Pour un état de jeu donné $(x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^m$ et un temps $t \in [0,T]$, on définit les processus :*

$$\begin{cases} (a^\star)_{t\le s\le T} := \left(\tilde{a}\left(s, X_s^{1,t,x_1}, X_s^{2,t,x_2}\right)\right)_{t\le s\le T}, \\ (b^\star)_{t\le s\le T} := \left(\tilde{b}\left(s, X_s^{1,t,x_1}, X_s^{2,t,x_2}\right)\right)_{t\le s\le T}. \end{cases}$$

Les processus $a^\star$ et $b^\star$ correspondent à des stratégies où chaque joueur identifie le champ de bataille ayant le plus d'impact sur la condition terminale espérée, et concentre ensuite tous ses efforts sur ce champ. On se retrouve alors dans le jeu avec une configuration où, à chaque instant t, il n'y a qu'un seul champ de bataille d'intérêt, où les deux joueurs consacrent tout leur budget.

Théorème 1.4.22. *Le couple de stratégies $(a^\star, b^\star)$ constitue un équilibre de Nash du jeu.*

Remarque 1.4.23. *En réalité, le couple $(a^\star, b^\star)$ possède une propriété plus forte : il forme un* point selle *du jeu.*

Théorème 1.4.24. *La valeur du jeu est donnée par w au sens suivant :*

$$J\big(t, x_1, x_2, a^\star, b^\star\big) \;=\; w\big(t, x_1 - x_2\big), \tag{1.19}$$

et, en particulier, $\underline{V}(t,x_1,x_2) = \overline{V}(t,x_1,x_2) = w\big(t, x_1 - x_2\big)$.

La preuve repose sur le principe de programmation dynamique et l'EDP HJBI correspondante. Dans le cas symétrique, le terme hamiltonien s'annule et l'EDP se réduit à l'équation de la chaleur rétrograde pour la variable $y = x_1 - x_2$, dont la solution est w. La vérification par Itô avec les stratégies $(a^\star, b^\star)$ donne (1.19). Voir le chapitre 2 pour les détails.

Cas particulier : le gain classique du jeu de Blotto

On présente dans cette partie un cas particulier de fonction de gain qui permet de formuler des résultats explicites, où la fonction de gain est donnée par

$$g(z) \;=\; \sum_{i=1}^{m} \lambda_i \mathbf{1}_{\{z_i > 0\}}, \qquad z \in \mathbb{R}^m,$$

où $(\lambda_i)_{i=1,\dots,m}$ sont des réels positifs. Il s'agit de la fonction de gain du jeu de Colonel Blotto standard. La fonction g est bornée et vérifie les hypothèses de régularité supposées, puisque son

ensemble de discontinuité est l'union des hyperplans $\{x_i = 0\}$, négligeables pour la mesure de Lebesgue.

Dans ce cas particulier, on peut calculer explicitement la fonction w. On obtient, pour tout $(t,x) \in [0,T) \times \mathbb{R}^m$:

$$w(t,z) \;=\; \sum_{i=1}^{m} \lambda_i \, \Phi\left(\frac{1}{\sqrt{2(T-t)}} \, z_i \right),$$

où Φ désigne la fonction de répartition de la loi normale standard.

On en déduit les dérivées partielles suivantes :

$$\partial_{x_i} w(t,z) \;=\; \frac{\lambda_i}{\sqrt{4\pi(T-t)}} \exp\left(-\frac{z_i^2}{4(T-t)} \right), \qquad i = 1, \dots, m.$$

De manière générale, la règle de passage d'un champ de bataille i à un autre j peut être caractérisée par la condition suivante :

$$\partial_{z_i} w(t,z) \;\leq\; \partial_{z_j} w(t,z) \;\iff\; z_i^2 - z_j^2 \;\geq\; 4(T-t) \, \log\left(\frac{\lambda_i}{\lambda_j} \right).$$

Cette écriture permet de faire quelques remarques sur le basculement entre champs de bataille : lorsque tous les coefficients λ_i sont égaux, chaque champ de bataille a le même poids. Dans ce cas, les deux joueurs concentrent la totalité de leur budget sur le champ où la bataille est la plus serrée, ce qui définit un équilibre de Nash.

Sinon, lorsque $t \ll T$, ce sont les coefficients λ_i qui dominent dans les choix stratégiques, tant que les différences entre les états des joueurs ne sont pas trop marquées. À l'inverse, quand $t \to T$, les différences $|z_i|$ prennent une place plus importante que les coefficients dans la décision stratégique : les joueurs privilégient alors les champs de bataille où l'écart entre les états est le plus faible.

Cas particulier : Application à un gain électoral

On considère le cas où, pour tout $z \in \mathbb{R}^m$, la fonction g est définie par

$$g(z) = \mathbf{1}_{\mathbb{R}_+} \left(\sum_{i=1}^{m} \mathbf{1}_{\mathbb{R}_+}(z_i) - \frac{m}{2} \right)$$

Ce modèle correspond par exemple aux élections présidentielles américaines, où les champs de bataille représentent les États. Et on gagne les élections si on gagne la majorité (plus que la moitié alors) des états. La fonction g satisfait les hypothèses nécessaires : elle est bornée, donc sa croissance est au plus exponentielle, et son ensemble de discontinuité

$$\left\{ z \in \mathbb{R}^m ; \; \sum_{i=1}^{m} \mathbf{1}_{\mathbb{R}_+}(z_i) = \frac{m}{2} \right\}$$

a une mesure de Lebesgue nulle.
On peut également caractériser explicitement la fonction w. Pour tout $(t,z) \in [0,T) \times \mathbb{R}^m$, on a

$$w(t,z) = \sum_{j=\lceil \frac{m}{2} \rceil}^{m} \left(\sum_{p_1 < p_2 < \ldots < p_j \in [\![1,m]\!]} \prod_{k=1}^{j} \Phi\left(\sqrt{\frac{1}{2(T-t)}} z_{p_k} \right) \right)$$

De plus, les dérivées partielles s'écrivent, pour tout $i \in [1,m]$ et $(t,x) \in [0,T) \times \mathbb{R}^m$, sous la forme

$$\partial_i w(t,z) = \frac{\exp\left[-\frac{z_i^2}{4(T-t)}\right]}{\sqrt{4\pi(T-t)}} \sum_{j=\lceil \frac{m}{2} \rceil}^{m} \left(\sum_{p_1<p_2<\ldots<p_{j-1} \in [\![1,m]\!]} \prod_{k=1}^{j-1} \Phi\left(\sqrt{\frac{1}{2(T-t)}} z_{p_k} \right) \right)$$

On peut ainsi caractériser les points où les stratégies des joueurs basculent d'un champ de bataille à un autre, même si cette description est moins intuitive. En effet, la relation $\partial_t w(t,z) = \partial_t w(t,z)$ s'écrit

$$\begin{aligned} &\frac{\Phi}{\phi}\Big(\sqrt{\tfrac{1}{2(T-t)}}\, z_l\Big)\, \frac{\phi}{\Phi}\Big(\sqrt{\tfrac{1}{2(T-t)}}\, z_i\Big)\, K_1^{j,l}(z) \\ &\quad + \frac{\phi}{\Phi}\Big(\sqrt{\tfrac{1}{2(T-t)}}\, z_i\Big)\, K_2^{j,l}(z) \\ &= K_1^{j,l}(z) + \frac{1}{\Phi\Big(\sqrt{\tfrac{1}{2(T-t)}}\, z_i\Big)} K_2^{j,l}(z). \end{aligned} \tag{1.20}$$

pour tout $(t,z) \in [0,T) \times \mathbb{R}^m$ et tout $j = 1, \ldots, m$, où ϕ désigne la densité de la loi normale centrée réduite, et où les fonctions $K^{j,1}$ et $K^{j,2}$ sont respectivement données par

$$K_1^{j,l}(z) = \sum_{j=\lceil \frac{m}{2} \rceil}^{m} \left(\sum_{p_1<p_2<\ldots<p_{j-2} \in [\![1,m]\!]\setminus\{i,j\}} \prod_{k=1}^{j-2} \Phi\left(\sqrt{\frac{1}{2(T-t)}} z_{p_k} \right) \right)$$

et

$$K_2^{j,l}(z) = \sum_{j=\lceil \frac{m}{2} \rceil}^{m} \left(\sum_{p_1<p_2<\ldots<p_{j-1} \in [\![1,m]\!]\setminus\{i,j\}} \prod_{k=1}^{j-1} \Phi\left(\sqrt{\frac{1}{2(T-t)}} z_{p_k} \right) \right)$$

Cas particulier : Critère spectral et cybersécurité

Dans cette sous-section, on présente des exemples issus de la cyber-sécurité. On considère des situations de compétition sur des réseaux modélisés par des graphes, ainsi que des fonctions objectif dépendant du spectre des matrices de Laplace (qu'on définit un peu plus loin) associées à ces graphes.

On modélise un réseau donné comme un graphe non orienté $G = (V,E)$, où V est l'ensemble des sommets et E l'ensemble des arêtes, avec des cardinalités respectives données par deux entiers $n \geq 2$ et $m \geq 1$. Comme d'usage en théorie des graphes, E sera un ensemble de couples non ordonnés d'éléments de V. Pour simplifier sans perte de généralité, on suppose que $E = \{1, \ldots, m\}$. On fait l'hypothèse qu'il n'y a ni boucle ni arêtes multiples.

Pour $(x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^m$, on note $G_{x_1 - x_2}$ le graphe pondéré (V, E, w), où le vecteur de poids w est défini par $w_j := 1$ si $x_1(j) - x_2(j) > 0$ et $w_j := 0$ sinon. Autrement dit, $G_{x_1-x_2}$ est le sous-graphe couvrant de G (c-à-d qui inclut tous les sommets de G) dont l'ensemble d'arêtes est $\widetilde{E} \subset E$, défini par

$$\widetilde{E} = \{e_j,\ j \in J(x_1 - x_2)\},$$

où $J(x_1 - x_2)$ désigne l'ensemble des indices tels que $x_1 - x_2$ soit strictement positif.

On considère maintenant la matrice laplacienne L associée au graphe $G_{x_1-x_2}$, définie par

$$L = D - A,$$

où A est la matrice d'adjacence de $G_{x_1-x_2}$ et D sa matrice des degrés associée, c'est-à-dire la matrice diagonale $n \times n$ dont l'entrée diagonale $D_{i,i}$ est le degré du sommet i. La matrice L est symétrique, semi-définie positive, et admet 0 comme valeur propre.

Soient $\lambda_1 = 0 \leq \lambda_2 \leq \cdots \leq \lambda_n$ les n valeurs propres de L. On considère des fonctions spectrales, c'est-à-dire des fonctions ψ que l'on peut écrire sous la forme $\psi(\lambda_1, \lambda_2, \ldots, \lambda_n) = \phi(w)$. On peut alors définir le critère du jeu spectral : étant donné un graphe $G = (V, E)$, on définit, pour $v \in \mathbb{R}^m$, la fonction $g(v) := \psi(G_v)$. Autrement dit, dans le jeu où la récompense est donnée par g, chaque arête du graphe G joue le rôle d'un champ de bataille. Si le joueur 1, qui cherche à maximiser son critère spectral associé, gagne le champ j, alors l'arête e_j est conservée dans le graphe ; sinon, elle est supprimée.

Remarque 1.4.25. *Un exemple intéressant de critère spectral est la connectivité algébrique du graphe, définie par $\psi(\lambda_1, \ldots, \lambda_n) = \lambda_2$. Cette valeur est strictement positive si et seulement si le graphe est connexe, et plus elle est élevée, plus le graphe est "bien interconnecté", dans le sens où il est difficile de le séparer en composantes par la suppression d'arêtes ou de sommets (voir Fiedler [30]). Dans un contexte de cybersécurité où les réseaux informatiques sont modélisés par des graphes, on peut interpréter λ_2 comme une mesure de la robustesse structurelle du réseau face à une attaque : un hacker cherchera à réduire cette valeur en supprimant des arêtes afin de rompre la connectivité globale.*

Le joueur 1 choisit la façon dont il répartit sa dérive de ressources sur les arêtes du graphe, avec pour objectif de maximiser

$$\mathbb{E}^{P^{a,b}}\left[\psi\left(G_{X_T^{1,t,x_1} - X_T^{2,t,x_2}}\right)\right],$$

tandis que le joueur 2 cherche à minimiser cette même quantité.

Théorème 1.4.26. *La fonction valeur du jeu spectral est donnée par*

$$V(t, x_1, x_2) = \sum_{\widetilde{G} \in \mathcal{S}} \psi(\widetilde{G}) \Phi_{\widetilde{G}}(T - t, x_1 - x_2),$$

où S désigne l'ensemble des sous-graphes couvrants de G et où

$$\Phi_{\widetilde{G}}(s, v) := \prod_{j \in \widetilde{E}} \Phi\left(\frac{v_j}{\sqrt{2s}}\right) \prod_{j \in \widetilde{E}^c} \overline{\Phi}\left(\frac{v_j}{\sqrt{2s}}\right), \quad \textit{pour } s > 0 \textit{ et } v \in \mathbb{R},$$

avec $\overline{\Phi} := 1 - \Phi$.

On peut également calculer des dérivées partielles de $w(t, x) = V(t, x_1, x_2)$, qui s'écrivent

$$\partial_i w(t, x) = \sum_{\widetilde{G} \in \mathcal{S}} \frac{\psi(\widetilde{G})\,(-1)^{\mathbf{1}_{i \in \widetilde{E}^c}}}{\sqrt{2(T-t)}} \phi\left(\frac{x_i}{2(T-t)}\right) \prod_{j \in \widetilde{E} \setminus i} \Phi\left(\frac{x_j}{\sqrt{2(T-t)}}\right) \prod_{j \in \widetilde{E}^c \setminus i} \overline{\Phi}\left(\frac{x_j}{\sqrt{2(T-t)}}\right)$$

1.4.7 EDSR et jeux différentiels stochastiques

On aborde à présent l'approche des jeux stochastiques par équations différentielles stochastiques rétrogrades (EDSR), ou Backward Stochastic Differential Equations (BSDE) qui permet de contourner la nécessité de résoudre des EDP. Sous la condition d'Isaacs qu'on introduira en 1.23, on peut relier la solution de l'EDP ainsi que son gradient à un objet probabiliste, solution d'une EDSR. Cette approche est détaillée dans l'article d'Hamadène et Lepeltier [36], sur lequel on s'appuiera pour introduire les concepts et résultats de cette sous-section.

Introduction des EDSR

Définition 1.4.27. *Étant donnés un espace probabilisé filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$, un mouvement brownien m-dimensionnel W, une variable aléatoire terminale ξ, et une fonction mesurable $f : [0,T]\times\Omega\times\mathbb{R}\times\mathbb{R}^m \to \mathbb{R}$, appelée* générateur*, une* solution *de l'équation différentielle stochastique rétrograde (EDSR) sur $[0,T]$ est la donnée d'un couple de processus (Y,Z) adapté à la filtration $(\mathcal{F}_t)_{0\le t\le T}$ tel que*

$$Y_t \;=\; \xi + \int_t^T f(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s, \qquad 0 \le t \le T. \tag{1.21}$$

Les solutions (Y,Z) sont recherchées dans les espaces suivants :

$$\mathcal{S}^2 := \Big\{ Y \text{ progressivement mesurable, càdlàg } :\ \mathbb{E}\Big[\sup_{0\le t\le T} |Y_t|^2\Big] < \infty \Big\},$$

$$\mathcal{H}^2 := \Big\{ Z \text{ progressivement mesurable à valeurs dans } \mathbb{R}^m \ :\ \mathbb{E}\Big[\int_0^T |Z_t|^2\, dt\Big] < \infty \Big\}.$$

Remarque 1.4.28. *On utilise souvent aussi la notation infinitésimale de l'EDSR 1.21 :*

$$dY_t = -f(t, Y_t, Z_t)\, dt + Z_t\, dW_t, \qquad Y_T = \xi. \tag{1.22}$$

Théorème 1.4.29 (Pardoux-Peng, 1990)**.** *Supposons que :*

1. *La variable terminale $\xi \in \mathcal{L}^2(\Omega, \mathcal{F}_T, \mathbb{P})$, c'est-à-dire $\mathbb{E}[|\xi|^2] < \infty$.*
2. *Le générateur $f : [0,T]\times\Omega\times\mathbb{R}\times\mathbb{R}^m \to \mathbb{R}$ est $\mathcal{F}_t$-progressivement mesurable et Lipschitz en (y,z), uniformément en (t,ω).*

Alors il existe une unique solution $(Y,Z) \in \mathcal{S}^2 \times \mathcal{H}^2$ de l'EDSR.

Ce théorème a été introduit et démontré par Pardoux et Peng [63].

Remarque 1.4.30. *L'EDSR se lit en sens inverse du temps : une condition terminale ξ est prescrite à T, et l'équation détermine rétroactivement l'évolution de (Y,Z) vers le passé. Il est cependant essentiel de noter que, bien que définis à rebours, les processus $(Y_t, Z_t)_{0\le t\le T}$ sont adaptés à la filtration $(\mathcal{F}_t)$, donc mesurables par rapport à l'information disponible jusqu'au temps t. Autrement dit, à chaque instant t, Y_t et Z_t dépendent uniquement du passé (et non du futur).*

Définition 1.4.31. *Étant donnés un espace probabilisé filtré $(\Omega, \mathcal{F}, (\mathcal{F}_s)_{t\le s\le T}, \mathbb{P})$, un mouvement brownien m-dimensionnel W, une condition initiale $X_t = x$, une condition terminale $Y_T = g(X_T)$, et des coefficients mesurables μ, σ, f, g convenablement réguliers, on appelle* équation différentielle stochastique progressive–rétrograde *(EDSPR), ou* Forward Backward Stochastic

Differential Equation *(FBSDE) sur* $[t,T]$*, un triplet de processus* (X,Y,Z) *adapté à la filtration* $(\mathcal{F}_s)_{t\le s\le T}$ *tel que, pour tout* $s\in[t,T]$*,*

$$\begin{cases} X_s \;=\; x+\int_t^s \mu(u,X_u,a_u,b_u)\,du+\int_t^s \sigma(u,X_u,a_u,b_u)\,dW_u,\\ Y_s \;=\; g(X_T)+\int_s^T f(u,X_u,Y_u,Z_u,a_u,b_u)\,du-\int_s^T Z_u\,dW_u, \end{cases}$$

soit vérifié.

Remarque 1.4.32. *Dans la formulation générale, les équations forward et backward sont* couplées *: les dynamiques de* X *peuvent dépendre de* (Y,Z) *et réciproquement. Dans ce manuscrit, on se restreint au cas où la dépendance se fait uniquement dans le sens* $X\mapsto(Y,Z)$*, c'est-à-dire que* X *n'est pas influencé directement par* Y *ou* Z*.*

Théorème 1.4.33 (Ma, 1994)**.** *On suppose que :*

1. *La condition initiale est donnée par* $X_t=x\in\mathbb{R}^d$*, et la condition terminale est de la forme* $Y_T=g(X_T)$*, où* $g:\mathbb{R}^d\to\mathbb{R}$ *est Lipschitz.*
2. *La dérive, le coefficient de diffusion, et le générateur* $\mu:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$*,* $\sigma:[0,T]\times\mathbb{R}^d\to\mathbb{R}^{d\times m}$ *et* $f:[0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^m\to\mathbb{R}$ *sont* $\mathcal{F}_s$*-progressivement mesurables et Lipschitz en* (x,y,z)*, uniformément en* (s,ω)*.*

Alors il existe une unique solution $(X,Y,Z)\in\mathcal{S}^2\times\mathcal{S}^2\times\mathcal{H}^2$ *du système EDSPR.*

Ce résultat est établi et rigoureusement démontré par Ma [57].

Application aux jeux

On considère la dynamique contrôlée suivante dans un jeu à somme nulle :

$$dX_t \;=\; \mu(t,X_t,a_t,b_t)\,dt \;+\; \sigma(t,X_t)\,dW_t,\qquad X_0=x,$$

et un critère

$$J(a,b) \;=\; \mathbb{E}\Big[g(X_T)\;+\;\int_0^T r(t,X_t,a_t,b_t)\,dt\Big],$$

que le Joueur 1 maximise en contrôlant a et le Joueur 2 minimise en contrôlant b.

Définition 1.4.34. *On définit le hamiltonien* h *suivant :*

$$h(t,x,z,a,b):=r(t,x,a,b)\;+\;\langle\sigma^{-1}\mu(t,x,a,b),\,z\rangle$$

La condition d'Isaacs, qu'on supposera vérifiée, correspond à

$$\inf_{b\in B}\sup_{a\in A}h(t,x,z,a,b)=\sup_{a\in A}\inf_{b\in B}h(t,x,z,a,b) \tag{1.23}$$

pour tout $(t,x,z)\in[0,T)\times\mathbb{R}^m\times\mathbb{R}^m$*.*

Définition 1.4.35. *Sous la condition d'Isaacs, on associe au jeu l'EDSR suivante :*

$$\begin{cases} Y_T \;=\; g(X_T),\\ dY_t \;=\; -\sup_{a\in A}\inf_{b\in B}\,h(t,X_t,Z_t,a,b)\,dt\;+\;Z_t\,dW_t,\qquad 0\le t\le T, \end{cases}$$

où la solution (Y,Z) *vit dans les espaces* $\mathcal{S}^2\times\mathcal{H}^2$ *définis plus haut.*

Remarque 1.4.36. *Dans le cas des jeux différentiels stochastiques à somme nulle, le générateur f de la partie backward est pris comme l'Hamiltonien h issu de la condition d'Isaacs. Ainsi, la solution (Y, Z) de l'EDSPR correspond directement à la fonction valeur du jeu (Y) et aux stratégies optimales associées (Z), ce qui relie la formulation probabiliste aux équations de Hamilton–Jacobi–Bellman–Isaacs.*

Théorème 1.4.37 (Hamadène–Lepeltier, 1995)**.** *Sous des hypothèses standard de Lipschitz et de croissance linéaire, ainsi que la condition d'Isaacs, l'EDSPR ci-dessus admet une solution unique $(Y^{t,x}, Z^{t,x})$. La fonction valeur du jeu en (t, x) est donnée par*

$$V(t,x) \;=\; Y_t^{t,x},$$

et il existe des stratégies markoviennes $(a^\star, b^\star)$ telles que, pour presque tout (t, x, z),

$$a^\star(t,x,z) \in \arg\max_{a\in A} \inf_{b\in B} h(t,x,z,a,b), \qquad b^\star(t,x,z) \in \arg\min_{b\in B} \sup_{a\in A} h(t,x,z,a,b).$$

En particulier, les stratégies markoviennes

$$\big(a^\star(t, X_t, Z_t),\, b^\star(t, X_t, Z_t)\big)$$

forment un point selle et réalisent la valeur V.

Ce théorème est démontré rigoureusement dans l'article d'Hamadène et Lepeltier [36].

Remarque 1.4.38. *L'un des intérêts principaux de l'approche par EDSR est qu'elle ne requiert aucune régularité a priori de la fonction valeur. En effet, si l'on pose $u(t,x) := Y_t^{t,x}$ et que l'on identifie $Z_t^{t,x}$ avec le gradient $\nabla_x u(t, X_t^{t,x})$ lorsque celui-ci existe, alors u est caractérisée comme une solution de viscosité de l'équation de Hamilton–Jacobi–Bellman–Isaacs correspondante. Ainsi, la formulation en termes d'EDSR fournit une représentation probabiliste de la fonction valeur et assure, par construction, son rôle de solution de viscosité de l'EDP associée.*

1.4.8 Contribution : jeu de Blotto différentiel stochastique asymétrique

On reprend le même formalisme que ce qu'on trouve dans 1.4.6, en prenant cette fois des budgets différents : $\alpha \neq \beta$.

Définition 1.4.39. *On introduit la fonction hamiltonienne $H : \mathbb{R}^m \times A \times B \to \mathbb{R}$ définie par*

$$H(z,a,b) := (a-b)\cdot z.$$

Définition 1.4.40. *On introduit les contrôles suivants, pour tout $z \in \mathbb{R}^m$:*

$$\tilde{a}(z) := \begin{cases} \alpha e_i & \text{si } \max_{j=1,\dots,m} z_j = z_i \text{ et } \max_{j=1,\dots,m} z_j > 0, \\ 0_{\mathbb{R}^m} & \text{sinon,} \end{cases}$$

et

$$\tilde{b}(z) := \begin{cases} \beta e_i & \text{si } \min_{j=1,\dots,m} -z_j = -z_i \text{ et } \min_{j=1,\dots,m} -z_j < 0, \\ 0_{\mathbb{R}^m} & \text{sinon,} \end{cases}$$

où, pour tout $i = 1, \dots, m$, $e_i \in \mathbb{R}^m$ désigne le i-ième vecteur de la base canonique de $\mathbb{R}^m$, et z_i est la i-ième coordonnée du vecteur z.

Lemme 1.4.41. *Les processus $\tilde{a}$ et $\tilde{b}$ vérifient, pour tout $(z,a,b) \in \mathbb{R}^m \times A \times B$:*

$$H\big(z,a,\tilde{b}(z)\big) \;\le\; H\big(z,\tilde{a}(z),\tilde{b}(z)\big) \;\le\; H\big(z,\tilde{a}(z),b\big).$$

On a alors l'identité suivante :

$$H\big(z,\tilde{a}(z),\tilde{b}(z)\big) \;=\; \inf_{b\in B}\sup_{a\in A} H(z,a,b) \;=\; \sup_{a\in A}\inf_{b\in B} H(z,a,b).$$

La preuve, détaillée dans le chapitre 2, repose sur la définition de $\tilde{a}$ et $\tilde{b}$. et la linéarité en a et b de H.

Lemme 1.4.42. *Soient $(a,b) \in \mathcal{A}\times\mathcal{B}$ et $(t,x_1,x_2) \in [0,T]\times\mathbb{R}^m\times\mathbb{R}^m$. Et soit $(Y^{a,b},Z^{a,b})$ l'unique solution de l'EDSR*

$$\begin{cases} Y_T^{a,b} = g\Big(X_T^{1,t,x_1} - X_T^{2,t,x_2}\Big), \\ dY_s^{a,b} = -(a_s-b_s)\cdot Z_s^{a,b}\,ds + Z_s^{a,b}\,dW_s, \qquad 0\le s\le T, \end{cases}$$

alors on a

$$J(t,x_1,x_2,a,b) = \mathbb{E}^{P^{a,b}}\Big[g\Big(X_T^{1,t,x_1} - X_T^{2,t,x_2}\Big)\;\big|\,\mathcal{F}_t\Big] = Y_t^{a,b}.$$

Lemme 1.4.43. *Soit $(\tilde{Y},\tilde{Z})$ la solution de l'EDSR de générateur $z \mapsto z\cdot\big(\tilde{a}(z)-\tilde{b}(z)\big)$, et de condition terminale $\tilde{Y}_T = g\Big(X_T^{1,t,x_1} - X_T^{2,t,x_2}\Big)$, $\quad (x_1,x_2)\in\mathbb{R}^m$, $0\le t\le s\le T$.*
On définit alors les contrôles markoviens

$$(a_s^\star, b_s^\star) \;:=\; \big(\,\tilde{a}(\tilde{Z}_s),\;\tilde{b}(\tilde{Z}_s)\,\big),$$

et l'on obtient

$$J(t,x_1,x_2,a^\star,b^\star) \;=\; \tilde{Y}_t,$$

c'est-à-dire que (a^*,b^*) forme un point selle du jeu.

La démonstration de ce résultat est détaillée dans le chapitre 2.

Remarque 1.4.44. *Le passage intéressant dans ce résultat repose sur l'utilisation $\tilde{a}$ et $\tilde{b}$ sur des ensembles d'optimisation définis par les* $\arg\max/\arg\min$ *du hamiltonien. Ce point garantit que les stratégies markoviennes $(a^\star,b^\star)$ sont bien définies, adaptées à la filtration, et utilisables dans le jeu stochastique.*

Le résultat du lemme 1.4.42 repose sur une caractérisation de l'équilibre en fonction de la solution à une EDSR. Pour pouvoir en tirer profit, on aborde brièvement dans la partie suivante la résolution numérique des EDPs.

1.4.9 Numérique

Les schémas aux *différences finies*, *volumes finis*, *éléments finis* et variantes *spectrales*/*DG* (Discontinuous Galerkin) pour les EDP elliptiques/paraboliques sont très utilisés en petites dimensions, et pour de bonnes raisons : ils offrent des garanties de consistance, de stabilité et de convergence. Leur principal désavantage réside dans le coût mémoire et temps de calcul typiquement exponentiel en la dimension m à cause du maillage de l'espace (et du temps) : c'est la *malédiction de la dimensionnalité*.

Cas des EDP linéaires

Pour des EDP *linéaires* de la forme $\partial_t u + \mathcal{L}u + c = 0$, la représentation de *Feynman–Kac* donne

$$u(t,x) = \mathbb{E}\Big[g(X_T) + \int_t^T c(s, X_s)\, ds \;\Big|\; X_t = x\Big],$$

où X suit l'EDS associée à $\mathcal{L}$. On peut alors estimer u par *Monte Carlo* en utilisant un schéma d'Euler–Maruyama sans grille spatiale, avec des variantes efficaces comme le *Multi-Level Monte Carlo* (MLMC) et des techniques classiques de réduction de variance. Cette approche contourne largement la malédiction pour des dimensions élevées.

EDP semi-linéaires et réseaux de neurones

Dès que l'EDP est *semi-linéaire* et dépend de u ou, surtout, de son gradient, la formule de Feynman–Kac standard ne s'applique plus. Quelques cas particuliers admettent encore une représentation *forward*, mais le cadre général requiert une formulation en *EDSR* :

$$dY_t = -f(t, X_t, Y_t, Z_t)\, dt + Z_t\, dW_t, \qquad Y_T = g(X_T).$$

On renvoie à E *et al.* [37] pour un recensement des méthodes utilisées pour résoudre numériquement les EDSR/EDP non linéaires. On s'intéresse ici à la méthode *Deep BSDE* [26] s'appuyant sur des réseaux de neurones profonds, visant précisément le régime haute dimension. Cette approche reformule l'EDSR sous forme de propagation *forward* contrainte par la condition terminale : on simule X en avant, et on minimise une perte terminale $\mathbb{E}\|Y_T - g(X_T)\|^2$, ce qui reconstruit numériquement la solution backward.

Cette approche n'utilise pas de maillage, exploite bien le parallélisme des réseaux de neurones, et se prête à des dimensions $m \geq 10$ quand les schémas sur mailles deviennent inabordables.

Dans le cadre du jeu de Blotto, on fait face à un problème semi-linéaire dépendant de z ; on adopte alors une stratégie Deep BSDE avec :

(i) paramétrisation directe de $Z(t,x)$ par réseau,

(ii) perte terminale complétée par des *termes locaux* (résiduel EDSR et cible de type Malliavin pour Z) afin d'identifier correctement Z le long de la trajectoire,

(iii) une relaxation (augmentation graduelle) sur le paramètre de non-linéarité pour stabiliser l'optimisation en grande dimension.

Schéma Deep BSDE modifié

On considère l'EDSPR markovienne

$$\begin{aligned} dX_t &= b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t, \qquad X_0 = x, \\ dY_t &= -f\big(t, X_t, Y_t, Z_t\big)\, dt + Z_t\, dW_t, \qquad Y_T = g(X_T), \end{aligned}$$

associée à l'EDP semi-linéaire

$$\partial_t u + \mathcal{L}u + f\Big(t, x, u, \sigma^\top(t,x)\nabla_x u\Big) = 0, \qquad u(T,x) = g(x),$$

où $\mathcal{L}$ est le générateur de X.

Pour approcher numériquement ce type d'EDSPR, on utilise un schéma de type *Deep BSDE*, inspiré de celui proposé par E–Han–Jentzen [26], que l'on adapte légèrement à notre cadre.

On fixe une grille temporelle $t_0 = 0 < t_1 < \cdots < t_N = T$. Étant donné un point initial $x \in \mathbb{R}^m$ et une EDSR de générateur f et de condition terminale g, on introduit :

— un paramètre scalaire apprenable $y_0 \in \mathbb{R}$, qui jouera le rôle d'approximation de Y_0 et que l'on déterminera par minimisation d'une fonction de perte,

— un réseau de neurones $\psi_Z : [0,T) \times \mathbb{R}^m \times \mathbb{R}^{d_Z} \to \mathbb{R}^m$, noté $(t, x, \theta_Z) \mapsto \psi_Z(t, x, \theta_Z)$, destiné à approximer le processus Z lorsque $X_t = x$.

Simulation des trajectoires On commence par simuler M trajectoires de X au moyen d'un schéma d'Euler–Maruyama, en utilisant des trajectoires simulées du mouvement brownien W. On note X_i^l et W_i^l la l-ième trajectoire simulée de X et de W au temps t_i ($1 \le l \le M$). On pose

$$\Delta W_i^l := W_{i+1}^l - W_i^l.$$

À partir de la condition initiale numérique $Y_0 := y_0$, on propage la solution le long de chaque trajectoire en utilisant la discrétisation de l'EDSR :

$$Y_{i+1}^l = Y_i^l - f\big(t_i, X_i^l, Y_i^l, \psi_Z(t_i, X_i^l, \theta_Z)\big)\,\Delta t_i + \psi_Z(t_i, X_i^l, \theta_Z)\,\Delta W_i^l, \qquad 0 \le i \le N-1.$$

Fonction de perte Pour simplifier les notations, on pose

$$Z_i^l := \psi_Z(t_i, X_i^l, \theta_Z), \qquad f_i^l := f\big(t_i, X_i^l, Y_i^l, Z_i^l\big).$$

On définit ensuite, pour chaque pas de temps et chaque trajectoire, le résidu local

$$r_i^l = Y_{i+1}^l - Y_i^l + f_i^l\,\Delta t_i - Z_i^l\,\Delta W_i^l,$$

ainsi qu'une cible de type Malliavin

$$\widehat{Z}_i^l = \frac{1}{\Delta t_i}\left(Y_{i+1}^l - Y_i^l + f_i^l\,\Delta t_i\right)(\Delta W_i^l)^\top.$$

La perte empirique utilisée pour l'apprentissage des paramètres (y_0, θ_Z) est

$$\mathcal{L}(y_0, \theta_Z) = \mathcal{L}_T(y_0, \theta_Z) + \lambda_{\text{res}}\,\mathcal{L}_{\text{res}}(y_0, \theta_Z) + \lambda_Z\,\mathcal{L}_Z(\theta_Z),$$

où $\lambda_{\text{res}}, \lambda_Z > 0$ sont des pondérations, et

$$\mathcal{L}_T = \frac{1}{M}\sum_{l=1}^{M} \|Y_N^l - g(X_N^l)\|^2, \qquad \mathcal{L}_{\text{res}} = \frac{1}{M}\sum_{l=1}^{M}\frac{1}{N}\sum_{i=0}^{N-1} \|r_i^l\|^2,$$

$$\mathcal{L}_Z = \frac{1}{M}\sum_{l=1}^{M}\frac{1}{N}\sum_{i=0}^{N-1} \|Z_i^l - \widehat{Z}_i^l\|^2,$$

où $\|\cdot\|$ désigne la norme euclidienne (et, pour les matrices, la norme de Frobenius).

Entraînement / optimisation Les paramètres optimaux sont définis par

$$(y_0^*, \theta_Z^*) \in \arg\min_{(y_0,\theta_Z)\in\mathbb{R}\times\mathbb{R}^{d_Z}} \mathcal{L}(y_0, \theta_Z). \tag{1.24}$$

On effectue cette minimisation au moyen de l'algorithme Adam. En pratique, on peut faire décroître progressivement λ_{res} et λ_Z au cours de l'entraînement. L'initialisation des paramètres est décrite plus loin.

Relaxation du générateur On exploite la continuité des solutions d'EDSR par rapport au générateur, comme dans [49, Proposition 2.1]. Pour $\delta \in [0,1]$, on pose

$$f^{\delta}(t,x,y,z) := \delta\, f(t,x,y,z),$$

et on entraîne successivement pour

$$\delta_0 = 0 < \delta_1 < \cdots < \delta_K = 1,$$

en réutilisant à chaque étape les paramètres obtenus à l'étape précédente, et en raffinant au besoin la grille temporelle. Le cas $\delta = 0$ correspond à l'équation de la chaleur, ce qui fournit un point de départ simple ; la continuité en δ permet ensuite de se rapprocher progressivement de l'EDSR d'origine.

Warm start Pour l'étape initiale $\delta_0 = 0$, on peut utiliser directement la valeur théorique de Y_0 comme condition initiale y_0, et ne randomiser que les paramètres du réseau approchant Z. Les résultats numériques correspondants sont présentés dans la section suivante.

Remarque 1.4.45. *La fonction $\mathcal{L}_T$ correspond à la perte introduite dans [26], qui ne pénalise que l'erreur sur la condition terminale. On ajoute ici un terme de résidu local $\mathcal{L}_{\text{res}}$, inspiré de [67], pour mieux faire respecter l'EDSR le long des trajectoires, ainsi qu'une pénalité $\mathcal{L}_Z$ de type cible de Malliavin, destinée à améliorer l'approximation de Z, ce qui est crucial pour l'extraction des stratégies.*

Remarque 1.4.46. *Les deux pénalités supplémentaires jouent des rôles différents : $\mathcal{L}_{\text{res}}$ impose l'identité discrète de l'EDSR et agit sur Y et Z, tandis que $\mathcal{L}_Z$ est une régression qui ne porte que sur le réseau ψ_Z.*

Remarque 1.4.47. *La cible de Malliavin exacte au temps t_i s'écrit*

$$Z_i^{\star} = \frac{1}{\Delta t_i}\, \mathbb{E}\big[(Y_{i+1} - Y_i + f_i\, \Delta t_i)\, (\Delta W_i)^{\top} \,\big|\, \mathcal{F}_{t_i}\big].$$

En pratique, on utilise l'estimateur Monte Carlo $\widehat{Z}_i^{l}$ défini ci-dessus et on minimise une erreur quadratique moyenne. Par la propriété de projection en L^2 de l'espérance conditionnelle, le minimiseur de cette erreur retrouve $Z_i^{\star}$ sans qu'il soit nécessaire de calculer explicitement l'espérance conditionnelle à l'intérieur de $\widehat{Z}_i^{l}$.

Résultats numériques du schéma Deep BSDE modifié

On considère un exemple de l'équation de la chaleur en dimension $m = 10$, avec un gain de type Blotto donné par

$$g(x) = \sum_{i=1}^{m} \alpha_i\, \mathbf{1}_{\mathbb{R}^+}(x_i), \qquad \alpha_i = 2\, \mathbf{1}_{\{i=1\}} + \mathbf{1}_{\{2 \le i \le m\}}.$$

Pour approximer Z, on utilise un réseau de neurones comportant deux couches cachées de taille 400, avec une fonction d'activation SiLU.

Pour le warm start, on fixe $\lambda_{\text{res}} = 0$ et $\lambda_Z = 0.06$. En initialisant y_0 à la valeur théorique 5.5 pour $x_0 = 0_{\mathbb{R}^{10}}$, et en initialisant aléatoirement les paramètres du réseau associé à Z, on obtient, après 5000 itérations, les résultats d'entraînement suivants :

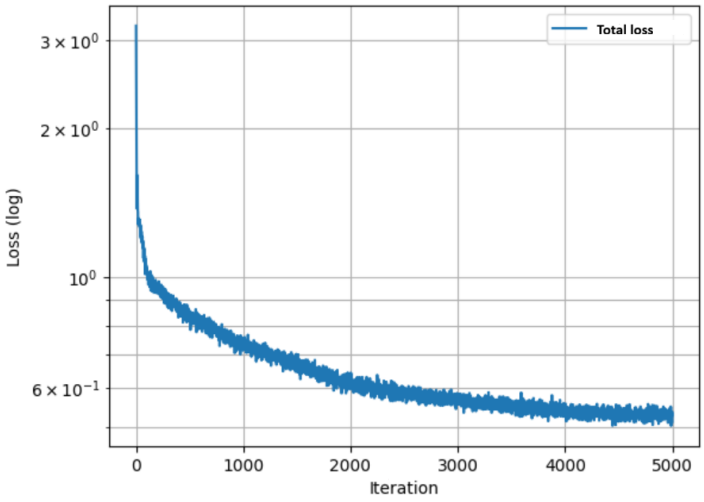


FIGURE 1.3 – Perte en fonction des itérations

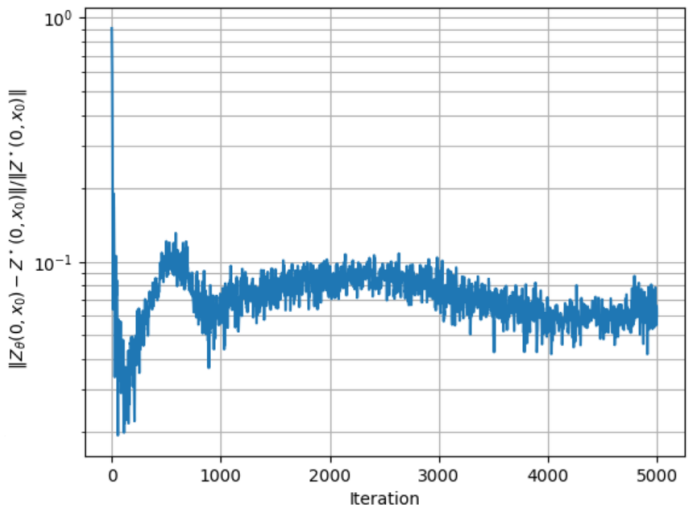


FIGURE 1.4 – Erreur relative sur Z en fonction des itérations

L'erreur relative finale sur Y est de 1.47×10^{-5}, et celle portant sur la norme de Z est de 6.005×10^{-2}.

On donne ensuite la courbe de gain pour le jeu de Blotto stochastique différentiel en fonction de la différence de budget

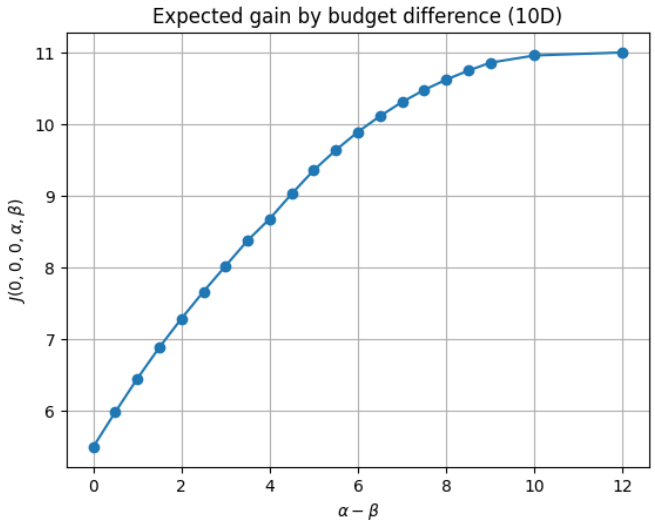


FIGURE 1.5 – Gain espéré du joueur 1 en fonction de la différence de budgets

Le code pour obtenir cette figure ainsi que le solver du schéma proposé sont disponibles en [40].

1.5 Théorie des contrats et jeux du Principal-Agent

La *théorie des contrats* s'intéresse aux situations où un Principal embauche un ou plusieurs agents pour effectuer une tâche, dans un contexte où leurs intérêts peuvent diverger et où certaines informations ne sont pas observables. L'enjeu central est alors de concevoir des contrats permettant d'aligner les incitations dans un cadre d'analyse rigoureux.

On se place alors dans le cadre d'un jeu de Stackelberg, c'est-à-dire une situation séquentielle où un joueur (le leader) choisit d'abord une action, et l'autre joueur (le follower) réagit ensuite en choisissant sa propre stratégie en réponse. Le leader anticipe donc la réaction du follower et intègre cette anticipation dans son choix initial. Ce cadre correspond de manière naturelle à la relation Principal–Agent : le Principal propose un contrat en premier, en sachant que l'Agent choisira ensuite son niveau d'effort optimal compte tenu de ce contrat.

Depuis les années 2000, le problème Principal-Agent a pris une nouvelle dimension en mathématiques financières, notamment grâce aux travaux de Sannikov [70] et à la formalisation de Cvitanić, Possamaï et Touzi [22]. Leur approche de *programmation dynamique* montre que l'on peut reformuler les problèmes de contrats en tant que problèmes de contrôle stochastique standard.

L'article d'Hernández-Santibáñez [44], sur lequel on se base pour l'introduction du formalisme de cette section, a proposé une extension unifiée qui assouplit les hypothèses des jeux du type Principal-Agent, tout en généralisant le cadre à un problème à plusieurs agents, chacun contrôlant son processus. En utilisant des techniques de changements de mesure et de formulations faibles des problèmes de contrôle stochastique, il établit des résultats permettant de traiter simultanément les modèles classiques à un Agent et les variantes multi-agents. Ce cadre fournit ainsi la base théorique d'une contribution de cette thèse.

1.5.1 Jeux de Principal-Agent différentiels stochastiques

Formulation des jeux stochastiques différentiels Principal-Agent

On présente dans cette section le problème général du principal agent ainsi que les résultats importants d'existence d'équilibres, puis on traitera le cas du jeu de Blotto en version principal-agent. On se met dans cette section dans toute cette section dans le cadre où un principal donne des contrats à N agents jouant un jeu différentiel stochastique entre eux, pour les inciter à effectuer des actions qui lui seront lucratives. On se propose dans cette section alors de formaliser cette idée.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilités et $T > 0$ un horizon de temps fini. Et soit $I = \{1, ..., N\}$ l'ensemble des N agents du problème. Chaque agent $i \in I$ est responsable du contrôle d'un processus d'état de dimension m, noté X^i. On note W un $\mathbb{P}$-mouvement brownien de dimension m.

Pour chaque $i = 1, \ldots, N$, on introduit une fonction

$$\sigma^i : \Omega \times [0, T] \times \mathbb{R}^{mN} \longrightarrow \mathcal{M}^{m,m},$$

qui représente la matrice de covariance associée au joueur i.

Définition 1.5.1. *Le processus X^i associé au joueur i est défini par*

$$X_t^i = X_0^i + \int_0^t \sigma_s^i(X_s^1, \dots, X_s^N)\, dW_s, \quad t \in [0,T], \tag{1.1}$$

où $X_0^i \in \mathbb{R}^m$ est supposé donné.

On peut écrire la dynamique des processus de sortie sous forme vectorielle en posant

$$X := \left((X^1)^\top, \dots, (X^N)^\top\right)^\top \in \mathbb{R}^{mN},$$

et en introduisant la matrice par blocs $\Sigma_s(X) \in \mathcal{M}^{mN,m}$, dont chaque bloc de ligne $\Sigma_s^i(\cdot)$ correspond à $\sigma_s^i(\cdot)$. On obtient alors la dynamique vectorielle équivalente :

$$X_t = X_0 + \int_0^t \Sigma_s(X_s)\, dW_s, \quad t \in [0,T]. \tag{1.2}$$

Remarque 1.5.2. *Hernández-Santibáñez [44] propose un cadre plus général en autorisant des dynamiques avec sauts. Dans ce manuscrit, on se restreint au cas continu afin de se concentrer sur les mécanismes d'incitation pertinents par rapport à la contribution de la thèse, sans alourdir les aspects techniques.*

Formulation faible

Dans l'esprit de la théorie des contrats, on adopte une formulation faible en termes de problème de martingale, où les actions des agents modifient la loi des processus d'état à travers un changement de probabilité. L'avantage de cette approche est que l'on n'a pas à démontrer l'existence de solutions fortes aux EDS associées : il suffit de caractériser les lois des processus sous des mesures de probabilité adaptées.

Pour chaque $i \in I$, soit A_i l'ensemble à dimension finie des actions possibles de l'agent i. On note $A := \prod_{i=1}^N A_i$ et $A^{-i} := \prod_{j \neq i} A_j$. Et on considère la fonction de drift suivante :

$$\mu : \Omega \times [0,T] \times \mathbb{R}^{mN} \times A \longrightarrow \mathbb{R}^m$$

telle que, pour tout $(y,a) \in \mathbb{R}^{mN} \times A$, le processus $(\mu_t(y,a))_{0 \le t \le T}$ soit $\mathbb{F}$-prévisible, c-à-d mesurable par rapport à la tribu engendrée par les processus adaptés, continus à gauche et à trajectoires càdlàg.

Définition 1.5.3. *On dit qu'un processus α à valeurs dans A, $\mathbb{F}$-prévisible, est une* action jointe admissible *si le processus suivant est une (vraie) $\mathbb{P}$-martingale :*

$$M_t^\alpha := \mathcal{E}\left(\int_0^t \mu_s(X_s, \alpha_s) \cdot dW_s\right), \tag{1.3}$$

où $\mathcal{E}(\cdot)$ désigne l'exponentielle stochastique. On note $\mathcal{A}$ l'ensemble des actions jointes admissibles.

Pour une action jointe admissible α, le processus M^α est une $\mathbb{P}$-martingale strictement positive. On définit alors une probabilité $\mathbb{P}^\alpha$, absolument continue par rapport à $\mathbb{P}$, dont la dérivée de Radon-Nikodym est donnée par M_T^α :

$$\frac{d\mathbb{P}^\alpha}{d\mathbb{P}}\Big|_{\mathcal{F}_T} = M_T^\alpha.$$

Il découle du théorème de Girsanov que le processus

$$W_t^\alpha := W_t - \int_0^t \mu_s(X_s, \alpha_s)\, ds$$

est un $\mathbb{P}^\alpha$-mouvement brownien, et que X admet la dynamique

$$X_t = X_0 + \int_0^t \Sigma_s(X_s)\mu_s(X_s, \alpha_s)\, ds + \int_0^t \Sigma_s(X_s)\, dW_s^\alpha \quad t \in [0,T], \tag{1.4}$$

Remarque 1.5.4. *Dans le cas où les dynamiques incluraient des sauts, il faudrait introduire des mesures de comptage et leurs compensateurs, ce qui entraînerait une écriture plus lourde et des conditions techniques supplémentaires explicitées dans Hernández-Santibáñez [44].*

Le jeu entre les agents

Chaque agent $i \in I$ est engagé par le Principal pour influencer la distribution de son processus d'état ainsi que ceux des autres agents. Les agents sont en concurrence et jouent donc un jeu différentiel stochastique à N joueurs, pour lequel on cherchera des équilibres de Nash.

Définition 1.5.5. *On note*

$$\mathcal{R} := \{\xi :\ \mathcal{F}_T\text{-mesurable, à valeurs réelles}\}$$
$$\mathcal{X} := \{\chi = (\chi_t)_{t\in[0,T]} : \text{à valeurs dans } \mathbb{R}^+; \mathbb{F}\text{-prévisible}\}$$

L'ensemble $\mathcal{R}$ correspond aux paiements finaux , tandis que $\mathcal{X}$ regroupe les paiements intermédiaires (ou paiements continus dans le temps). Un contrat *pour l'agent $i \in I$ est un couple $(\xi^i, \chi^i) \in \mathcal{R} \times \mathcal{X}$; pour l'ensemble des agents, on écrira $(\xi, \chi) \in \mathcal{R}^N \times \mathcal{X}^N$.*

Définition 1.5.6. *Pour un contrat (ξ, χ) et une action jointe $\alpha \in \mathcal{A}$, on définit l'*utilité *de l'agent i par :*

$$U_0^i(\alpha^i, \alpha^{-i}; \xi^i, \chi^i) := \mathbb{E}^{\mathbb{P}^\alpha}\Bigg[e^{-\int_0^T \rho_s^i(X_s,\chi_s,\alpha_s)\, ds}\, \mathcal{U}_A^i(\xi^i) + \int_0^T e^{-\int_0^s \rho_u^i(X_u,\chi_u,\alpha_u)\, du}\Big(u_A^i(\chi_s^i) - c_s^i(X_s,\alpha_s)\Big)\, ds\Bigg],$$

où :

- $\mathcal{U}_A^i : \mathbb{R} \to \mathbb{R}$ *est une fonction d'utilité strictement croissante et continue,*
- $u_A^i : \mathbb{R}^N \to \mathbb{R}$ *est une fonction d'utilité mesurable,*
- $c^i : \Omega \times [0,T] \times \mathbb{R}^{mN} \times A \to \mathbb{R}_+$ *est le coût instantané des actions choisies,*
- $\rho^i : \Omega \times [0,T] \times \mathbb{R}^{mN} \times \mathbb{R}^N \times A \to \mathbb{R}$ *est le taux d'actualisation instantané,*
- $\mathbb{P}^\alpha$ *est la probabilité induite par le changement de mesure associé à l'action α.*

Remarque 1.5.7. *On note que $\mathcal{U}_A^i$ n'est pas forcément concave. En pratique, cependant, les fonctions d'utilité sont souvent supposées concaves afin de refléter le principe des rendements décroissants, c'est-à-dire que l'utilité marginale d'une unité supplémentaire de ressource ou d'effort diminue à mesure que le niveau total augmente. Cette hypothèse traduit une attitude de type aversion au risque et permet d'obtenir un cadre mathématique plus régulier.*

Remarque 1.5.8. *Quand le gain à l'instant terminal et celui que l'agent gagne à chaque instant s'écrivent de manière indépendante, on dit que l'utilité U_0^i est* séparable. *Cependant, s'il existe un facteur d'actualisation couplé commun, on dit que l'utilité U_0^i est* quasi-séparable. *En particulier, l'utilité est supposée dans la définition 1.5.6 quasi-séparable.*

On formalise à présent la notion de meilleure réponse, puis d'équilibre de Nash.

Définition 1.5.9. *La valeur de l'agent i correspondant à une réponse optimale aux actions α^{-i} des autres est donnée par*

$$V_0^i(\alpha^{-i};\xi^i,\chi^i) := \sup_{\alpha^i\in\mathcal{A}_i(\alpha^{-i})} \mathbb{E}^{\mathbb{P}^\alpha}\Bigg[e^{-\int_0^T \rho_s^i(X_s,\chi_s,\alpha_s)\,ds}\,\mathcal{U}_A^i(\xi^i) + \int_0^T e^{-\int_0^s \rho_u^i(X_u,\chi_u,\alpha_u)\,du}\Big(u_A^i(\chi_s^i) - c_s^i(X_s,\alpha_s)\Big)\,ds\Bigg], \tag{1.5}$$

où l'ensemble des contrôles admissibles pour i est

$$\mathcal{A}_i(\alpha^{-i}) := \Big\{\alpha^i \in \ A_i;\ \mathbb{F}\text{-}\textit{prévisible, tel que } \alpha^i\otimes_i\alpha^{-i}\in\mathcal{A}\Big\}.$$

où on utilise l'opérateur $\otimes_i$ pour exprimer la concaténation de l'action α^i du joueur i avec celles des autres joueurs α^{-i}, de sorte que $\alpha^i\otimes_i\alpha^{-i}$ représente une action jointe admissible.

Définition 1.5.10. *Pour un contrat $(\xi,\chi)\in\mathcal{R}^N\times\mathcal{X}^N$, une action jointe $\alpha^\star\in\mathcal{A}$ est un* équilibre de Nash *si, pour tout $i\in I$,*

$$V_0^i\big((\alpha^\star)^{-i};\xi^i,\chi^i\big) \ = \ U_0^i\big((\alpha^\star)^i,(\alpha^\star)^{-i};\xi^i,\chi^i\big).$$

On note $NE(\xi,\chi)$ l'ensemble des équilibres de Nash associés au contrat (ξ,χ).

On définit ensuite la valeur dynamique d'un agent.

Définition 1.5.11. *Étant donné un contrat $(\xi^i,\chi^i)\in\mathcal{R}\times\mathcal{X}$ pour l'agent i, et des actions α^{-i} des autres agents, on définit pour tout temps d'arrêt $\tau\in\mathcal{T}_{0,T}$ la* valeur dynamique *de l'agent i par*

$$V^i(\tau,\alpha^{-i};\xi^i,\chi^i) := \operatorname*{ess\,sup}_{\alpha^i\in\mathcal{A}_i(\alpha^{-i})} \mathbb{E}_\tau^{\mathbb{P}^\alpha}\Bigg[e^{-\int_\tau^T \rho_s^i(X_s,\chi_s,\alpha_s)\,ds}\,\mathcal{U}_A^i(\xi^i) + \int_\tau^T e^{-\int_\tau^s \rho_u^i(X_u,\chi_u,\alpha_u)\,du}\Big(u_A^i(\chi_s^i) - c_s^i(X_s,\alpha_s)\Big)\,ds\Bigg], \tag{1.6}$$

où $\mathcal{T}_{0,T}$ désigne l'ensemble des temps d'arrêt à valeurs dans $[0,T]$, et $\mathbb{E}_\tau^{\mathbb{P}}$ l'espérance conditionnelle sous $\mathbb{P}$ par rapport à $\mathcal{F}_\tau$.

Remarque 1.5.12. *Le processus $V^i(\tau,\alpha^{-i};\xi^i,\chi^i)$ représente la valeur de continuation de l'agent i au temps τ, c'est-à-dire le gain anticipé qu'il peut obtenir entre τ et T en choisissant de manière optimale sa stratégie, étant donné les actions des autres agents. Ce processus sera très utile plus tard dans les résultats d'existence d'équilibres.*

Contraintes de participation. Dans la suite, le Principal n'offre que des contrats (ξ, χ) qui satisfont deux conditions :

— **Existence d'un équilibre.** Le jeu induit par le contrat admet au moins un équilibre de Nash, c'est-à-dire $NE(\xi, \chi) \neq \emptyset$.

— **Acceptation des agents.** Il existe $\alpha^\star \in NE(\xi, \chi)$ tel que

$$U_0^i\big((\alpha^\star)^i, (\alpha^\star)^{-i}; \xi^i, \chi^i\big) \;\geq\; R_0^i, \qquad \forall i \in I,$$

où R_0^i désigne la valeur de réservation de l'agent i, autrement dit son utilité minimale pour accepter de participer au contrat (niveau exogène, supposé connu du Principal).

Le problème du principal

Le Principal engage les N agents et cherche à maximiser son utilité espérée, en anticipant les équilibres de Nash qui résultent des contrats proposés.

Définition 1.5.13. *Seuls les contrats dits* admissibles *sont considérés, c'est-à-dire ceux qui garantissent que*

(i) le problème de principal-agent est bien posé,

(ii) les agents acceptent le contrat,

(iii) il existe un équilibre de Nash entre les agents.

On note $\mathcal{C}$ l'ensemble des contrats admissibles.

Remarque 1.5.14. *Les conditions d'admissibilité sont en réalité plus techniques que ce qu'on vient d'énoncer. On se contentera dans ce manuscrit de cette formulation, et on renvoie à Hernández-Santibáñez [44] pour une version plus extensive.*

Définition 1.5.15. *Le problème du Principal s'écrit alors :*

$$V_P := \sup_{(\xi,\chi)\in\mathcal{C}} \sup_{\alpha^\star \in NE(\xi,\chi)} \mathbb{E}^{\mathbb{P}^{\alpha^\star}}\left[e^{-\int_0^T r_s(X_s)\,ds}\, \mathcal{U}_P\Big(L(X_T) - \sum_{i=1}^N \xi^i\Big) - \int_0^T e^{-\int_0^s r_u(X_u)\,du}\, u_P(\chi_s)\,ds \right]$$

où

— $\mathcal{U}_P : \mathbb{R}^N \to \mathbb{R}$ *mesurable, est une fonction d'utilité terminale,*

— $u_P : \mathbb{R}^N \to \mathbb{R}$ *mesurable, est une fonction d'utilité instantanée,*

— $L : \mathbb{R}^{mN} \to \mathbb{R}$ *est une fonction de liquidation qui associe un gain agrégé aux sorties X_T,*

— $r : \Omega \times [0,T] \times \mathbb{R}^{mN} \to \mathbb{R}_+$ *est un taux d'actualisation instantané, où $r_t(x)_{0\leq t\leq T}$ est $\mathbb{F}$-prévisible pour chaque $x \in \mathbb{R}^{mN}$.*

Remarque 1.5.16. *On rappelle que l'ensemble des équilibres de Nash n'est a priori pas réduit à un singleton, il est donc nécessaire de spécifier un critère de sélection. On suppose le plus souvent dans la littérature que le principal peut imposer l'équilibre de Nash qui maximise son utilité (d'où le* sup *sur NE). Cette hypothèse reflète le pouvoir de négociation élevé du principal. Cependant, dans un contexte où ce pouvoir serait limité, le problème serait une optimisation avec contrainte d'appartenance de l'équilibre de Nash à un sous-ensemble. Alternativement, on peut supposer qu'il existe une règle de sélection d'équilibres, connue du principal.*

On a à présent défini le jeu des agents ainsi que le problème de contrôle du principal. On peut désormais passer aux résultats d'existence de solutions et d'équilibres. On note cependant l'ordre de résolution : on commence par trouver un équilibre du jeu des agents à contrat donné. Puis on utilise ce résultat pour que le principal puisse anticiper cet équilibre afin de résoudre son propre problème, c'est-à-dire proposer aux agents le contrat le plus intéressant pour lui.

Résolution du jeu des agents

Soit (ξ, χ) un contrat qu'on supposera donné aux agent pour cette partie. Afin de caractériser les équilibres de Nash associés à ce contrat donné, on introduit la fonction suivante :

Définition 1.5.17. *On définit*

$$f : \Omega \times [0,T] \times \mathbb{R}^{mN} \times \mathbb{R}^N \times \mathcal{M}^{N,mN} \times \mathbb{R}^N \times A \longrightarrow \mathbb{R}^N,$$

composante par composante, pour tout $i \in I$, par

$$f_s^i(x,y,z,k,a) = u_A^i(k) - c_s^i(x,a) - \rho_s^i(x,k,a)\, y^i + z^i \cdot \Sigma_s(x)\mu_s(x,a).$$

Cette fonction permet alors de définir une stratégie jointe d'équilibre :

Définition 1.5.18. *Étant donné $(s,x,y,z,k) \in [0,T] \times \mathbb{R}^{mN} \times \mathbb{R}^N \times \mathcal{M}^{N,mN} \times \mathbb{R}^N$, on définit une* action jointe d'équilibre *$a^\star(s,x,y,z,k)$ comme tout élément de A tel que pour tout $i \in I$,*

$$a^{\star,i}(s,x,y,z,k) \in \arg\max_{a^i \in A_i} f_s^i\big(x,y,z,k, a^i \otimes_i (a^{\star,-i}(s,x,y,z,k))\big).$$

Hypothèse. *On suppose dans la suite que $a^\star$ existe et est unique.*

On s'intéresse à présent à f si on y injecte la stratégie jointe d'équilibre :

Définition 1.5.19. *On définit alors la fonction*

$$F_s(x,y,z,k) := f_s\big(x,y,z,k,a^\star(s,x,y,z,k)\big)$$

On considère alors l'équation différentielle stochastique rétrograde (EDSR) multidimensionnelle suivante :

$$Y_t = \mathcal{U}_A(\xi) + \int_t^T F_s(X_s,Y_s,Z_s,\chi_s)\, ds - \int_t^T Z_s\, dX_s, \qquad t \in [0,T], \tag{1.7}$$

où $\mathcal{U}_A(\xi) \in \mathbb{R}^N$ désigne le vecteur dont la i-ème coordonnée est $\mathcal{U}_A^i(\xi^i)$. On introduit à présent une nouvelle définition d'EDSR qui permettra d'alléger les conditions d'intégrabilité :

Définition 1.5.20. *Une solution de l'EDSR* (1.7) *est un couple (Y,Z) tel que*

- *Y est un processus $\mathbb{F}$-adapté, à valeurs dans $\mathbb{R}^N$.*
- *$Z \in \mathbb{H}^{2,N\times mN}_{\mathrm{loc}}(X)$, c'est-à-dire un processus $\mathbb{F}$-prévisible, à valeurs dans $\mathcal{M}^{N,mN}$, et admettant une suite de temps d'arrêts $(\tau_n)_n \underset{n\to\infty}{\longrightarrow} T$ telle que $(Z_{t\wedge\tau_n})_{0\le t\le T} \in \mathbb{H}^{2,N\times mN}(X)$, défini comme suit :*

$$\mathbb{H}^{2,N\times mN}(X) := \left\{ Z \text{ à valeurs dans } \mathcal{M}^{N,mN}, \mathbb{F}\text{-prévisible, avec } \mathbb{E}\left[\int_0^T |Z_s\Sigma_s(X_s)|^2\, ds\right] < \infty \right\}$$

Remarque 1.5.21. *Demander que Z soit dans $\mathbb{H}^{2,N\times mN}_{\mathrm{loc}}(X)$ allège la condition d'intégrabilité globale en une d'intégrabilité locale. En effet, c'est une manière de demander le minimum nécessaire pour que les intégrales stochastiques aient un sens.*

Hypothèse. On suppose par la suite que F est telle qu'il existe, pour tout $y \in \mathbb{R}^N$ et tout $Z \in \mathbb{H}^{2,N\times mN}_{\mathrm{loc}}(X)$, une unique solution forte à l'EDS :

$$Y_t^{y,Z,\chi} = y - \int_0^t F_s(X_s, Y_s^{y,Z,\chi}, Z_s, \chi_s)\, ds + \int_0^t Z_s\, dX_s^c, \qquad t \in [0,T]. \tag{1.8}$$

Il faut encore une dernière définition pour énoncer les résultats d'existence de cette partie :

Définition 1.5.22. *Soit $y \in \mathbb{R}^N$, on note $\mathcal{Z}(y,\chi) \subset \mathbb{H}^{2,N\times mN}_{\text{loc}}(X)$ l'ensemble des Z pour lesquels*

(i) l'EDS 1.8 admet une solution

(ii) $\Big(a^\star(s, X_s, Y_s^{y,Z,\chi}, Z_s, \chi_s)\Big)_s$ est admissible

(iii) les variables aléatoires $\int_0^T |Z_s\Sigma_s(X_s)|^2 ds$, $\int_0^T |u_P(\chi_s)|\, ds$, et $\mathcal{U}_P\Big(L(X_T) - \sum_i \xi^i\Big)$ sont intégrables, où $\xi^i := (\mathcal{U}_A^i)^{-1}\big(Y_T^{y,Z,\chi,i}\big)$

(iv) pour chaque i, $Y_T^{y,Z,\chi,i}$ appartient à l'image de $\mathcal{U}_A^i$, de sorte que ξ^i soit bien définie.

Proposition 1.5.23. *Toujours à (ξ,χ) un contrat admissible fixe, soit $a^\star \in NE(\xi,\chi)$ une action jointe d'équilibre. Alors, il existe un couple (Y,Z) solution de l'équation rétrograde stochastique (EDSR)*

$$Y_t \;=\; \mathcal{U}_A(\xi) + \int_t^T F_s(X_s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dX_s, \qquad t \in [0,T], \tag{1.9}$$

tel que l'action jointe s'écrive sous la forme

$$a_s^{\star,i} \;=\; a^{\star,i}(s, X_s, Y_s, Z_s), \qquad dt \otimes d\mathbb{P}\text{-p.s., pour tout } i \in I.$$

Réciproquement, si (Y,Z) est solution de cette EDSR, alors le contrôle défini par

$$a_s^{\star,i} \;=\; a^{\star,i}(s, X_s, Y_s, Z_s), \qquad i \in I,$$

appartient à $NE(\xi,\chi)$.

Remarque 1.5.24. *Ce résultat est fondamental : il établit une équivalence entre l'existence de stratégies d'équilibre et celle d'une solution de l'EDSR. La preuve, détaillée par Hernández-Santibáñez [44], repose sur des arguments liés à la structure du problème principal-agent.*

Corollaire 1.5.25. *Soit $y \in \mathbb{R}^N$, et $Z \in \mathcal{Z}(y,\chi)$. Soit pour chaque $i \in I$, $\xi^i = (\mathcal{U}_A^i)^{-1}(Y_T^{y,Z,\chi,i})$, où $Y^{y,Z,\chi}$ est solution de l'EDS 1.8. Alors il existe une action jointe $a^\star \in NE(\xi,\chi)$ telle que les valeurs des agents vérifient*

$$V_0^i\big((a^\star)^{-i}; \xi^i, \chi^i\big) \;=\; y^i, \qquad \forall i \in I.$$

Remarque 1.5.26. *Ce corollaire montre que, sans perte de généralité, il suffit de considérer les contrats de la forme $\xi^i = (\mathcal{U}_A^i)^{-1}(Y_T^{y,Z,\chi,i})$. Cette observation permet de relier le problème du principal à un cadre de contrôle stochastique, qui sera exploité dans la reformulation suivante.*

Reformulation et résolution du problème du principal

Avec les résultats qu'on vient d'établir, on peut reformuler le problème du principal comme un problème de contrôle stochastique classique :

Théorème 1.5.27. *On a*

$$V_P \;=\; \sup_{y \geq R_0} v_P(y),$$

avec

$$v_P(y) := \sup_{\chi \in \pi_\chi \mathcal{C}} \sup_{Z \in \mathcal{Z}(y,\chi)} \mathbb{E}^{\mathbb{P}^{a^\star}}\Bigg[e^{-\int_0^T r_s(X_s)\, ds}\, \mathcal{U}_P\Big(L(X_T) - \sum_{i=1}^N (\mathcal{U}_A^i)^{-1}(Y_T^{y,Z,\chi,i})\Big) - \int_0^T e^{-\int_0^s r_u(X_u)\, du}\, u_P(\chi_s)\, ds \Bigg],$$

où R_0 désigne le vecteur des valeurs de réservation des agents et $\pi_\chi \mathcal{C}$ l'ensemble des paiements instantanés admissibles, défini par projection de $\mathcal{C}$ sur sa deuxième composante.

On rappelle les dynamiques :

$$\begin{aligned} X_t &= X_0 + \int_0^t \Sigma_s(X_s)\,\mu_s\big(X_s, a_s^{\star,y,Z,\chi}\big)\,ds + \int_0^t \Sigma_s(X_s)\,dW_s^{a^\star,y,Z,\chi} \\ Y_t^{y,Z,\chi} &= y - \int_0^t F_s\big(X_s, Y_s^{y,Z,H,\chi}, Z_s, \chi_s\big)\,ds + \int_0^t Z_s\Sigma_s(X_s)\,\mu_s\big(X_s, a_s^{\star,y,Z,\chi}\big)\,ds \\ &\quad + \int_0^t Z_s\Sigma_s(X_s)\,dW_s^{a^\star,y,Z,\chi} \end{aligned}$$

Proposition 1.5.28. *Le problème du principal est un problème de contrôle stochastique auquel on peut associer l'équation HJB suivante :*

$$-\partial_t v(t,x,y) - H_t\big(x,y,v,D_x v, D_y v, D_{yy}^2 v, D_{xy}^2 v\big) - \tfrac{1}{2}\mathrm{Tr}\Big(\Sigma_t(x)\Sigma_t(x)^\top D_{xx}^2 v(t,x,y)\Big) = 0 \quad (1.10)$$

Pour tout $(t,x) \in [0,T) \times \mathbb{R}^{mN}$, avec condition terminale

$$v(T,x,y) \;=\; \mathcal{U}_P\left(L(x) - \sum_{i=1}^N \big(\mathcal{U}_A^i\big)^{-1}(y^i)\right), \quad (x,y) \in \mathbb{R}^{mN} \times \mathbb{R}^N.$$

Le Hamiltonien H_t associé est défini par

$$H_t(x,y,u,p,\tilde{p},\tilde{q},r) \;:=\; \sup_{(z,k)\in\mathcal{M}^{N,dN}\times\mathbb{R}^N} \phi_t(x,y,u,p,\tilde{p},\tilde{q},r,z,k,a^\star(t,x,y,z,k)),$$

où

$$\begin{aligned} \phi_t(x,y,u,p,\tilde{p},\tilde{q},r,z,k,a) := {} & -u_P(k) - r_t(x)u + p\cdot\Sigma_t(x)\mu_t(x,a) + \tilde{p}\cdot z\Sigma_t(x)\mu_t(x,a) \\ & - \tilde{p}\cdot F_t(x,y,z,k) + \tfrac{1}{2}\,\mathrm{Tr}\big(z\Sigma_t(x)\Sigma_t(x)^\top z^\top \tilde{q}\big) + \mathrm{Tr}\big(z\Sigma_t(x)\Sigma_t(x)^\top r\big) \end{aligned}$$

Ainsi, le problème du Principal se reformule entièrement comme un problème de contrôle stochastique classique, gouverné par l'équation de Hamilton–Jacobi–Bellman (1.10). Cette caractérisation fournit une description complète des équilibres du jeu et clôt la résolution du problème.

1.5.2 Contribution : Jeu P-A différentiel stochastique de Blotto pessimiste

On propose dans cette partie une extension du jeu différentiel stochastique de Blotto au cadre de la théorie des contrats. On suppose qu'un principal (par exemple une entreprise) embauche un agent (spécialiste de cybersécurité) pour qu'il joue à une variante du jeu de Blotto contre un tiers (hacker).

Le principal propose à l'agent un contrat composé d'un paiement terminal accompagné d'un budget maximal dynamique, coûteux pour le principal. Son objectif est de maximiser son utilité espérée en tenant compte des stratégies optimales adoptées par l'agent et par le hacker. L'interaction entre l'agent et le hacker est formalisée comme un jeu de type *Blotto stochastique dynamique*. Le problème du Principal est alors de concevoir le contrat de manière cohérente avec ce cadre de contrôle optimal.

Formulation du problème

Dans cette partie, on fixe un horizon de temps fini $T > 0$ ainsi qu'un nombre $m \in \mathbb{N}^\star$ de champs de bataille. Le système évolue sur l'espace canonique des trajectoires continues à valeurs dans $\mathbb{R}^m$, muni de sa tribu borélienne. On note $\mathbb{P}$ la mesure de Wiener sur $(\Omega, \mathcal{F})$ et $\mathbb{F}$ la filtration canonique associée. Pour $\lambda > 0$, on définit

$$A_\lambda := \big\{x \in \mathbb{R}^m,\ \forall i = 1, \dots, m,\ x_i \geq 0,\ \sum_{i=1}^m x_i \leq \lambda\big\}.$$

On pose $B := A_\beta$ où β est une constante positive représentant le budget du hacker, fixé pour la suite et supposé connu publiquement. On note $\mathcal{A}$ l'ensemble des processus à valeurs réelles positives et $\mathbb{F}$-adaptés. On note aussi $\mathcal{B}$ l'ensemble des processus $\mathbb{F}$-progressivement mesurables à valeurs dans B. Et, pour tout processus $\alpha = (\alpha_t)_{0 \leq t \leq T} \in \mathcal{A}$, on note $\mathcal{A}(\alpha)$ l'ensemble des processus $\mathbb{F}$-adaptés y tels que $y_s \in A_{\alpha_s}$ pour tout $0 \leq s \leq T$. Ces deux ensembles représenteront alors les espaces de contrôles admissibles des joueurs.

Le jeu est écrit en formulation faible (cf. Définition 1.5.3). Pour un budget $\alpha \in \mathcal{A}$ donné, et profil de contrôle $(a, b) \in \mathcal{A}(\alpha) \times \mathcal{B}$, on aura le processus d'état :

$$X_s^{t,x} := x + \int_t^s (a_u - b_u)du + \int_t^s W_u^{a,b} du, \quad s \in [t, T]. \tag{1.11}$$

Avec le changement de probabilités :

$$\frac{dP^{a,b}}{dP} := \exp\left(\int_0^T (a_s - b_s)\mathrm{d}X_s - \frac{1}{2}\int_0^T |a_s - b_s|^2 \mathrm{d}s\right).$$

où on aura que le processus $(W_u^{a,b})_u$ est un mouvement Brownien de dimension m sous $P^{a,b}$.

Pour alléger les notations, on omettra de mentionner les termes (t, x) pour la suite du chapitre, et on les spécifiera si besoin.

Le Principal choisit un budget $\alpha \in \mathcal{A}$, où $\mathcal{A}$. Ce budget représente les ressources dont dispose dynamiquement l'Agent pour mener ses actions sur les champs de bataille ainsi qu'une possibilité d'arbitrage sur son propre effort. En plus, le Principal propose à l'Agent un contrat $\xi_A \in \Xi$, où l'ensemble Ξ est constitué de variables aléatoires $\mathcal{F}_T^X$-mesurables admettant des moments exponentiels de tout ordre. Autrement dit, le Principal rémunère l'Agent à l'échéance T, en fonction de la trajectoire du processus d'état.

Étant donné un contrat ξ_A et un processus de budget α, l'Agent reçoit la rémunération ξ_A au temps T. Pendant la compétition, i.e. à chaque instant entre 0 et T, il peut également détourner une partie du budget qui lui est alloué afin de réaliser un profit supplémentaire, son action $a \in \mathcal{A}(\alpha)$ n'étant pas observable par le Principal. On suppose que l'Agent a une utilité exponentielle.

Enfin, pour le dernier participant, on supposera que le Hacker a pour objectif de saboter une caractéristique du Principal. Cela est modélisé par un Principal externe (nommé Principal-hacker) qui offrira au hacker un contrat ξ_H qui modélisera l'objectif **à minimiser** de son intervention. Le jeu entre l'Agent et le Hacker consiste alors à ce que l'Agent maximise son critère de jeu, défini ci-dessous, tandis que le Hacker cherche à minimiser son critère via ξ_H.

Définition 1.5.29. *Étant donné un contrat ξ_A, un budget α choisi par le Principal et un contrat ξ_H offert au Hacker, on définit le critère de jeu par*

$$J_A(\xi, \alpha, a, b) := \mathbb{E}^{\mathbb{P}^{a,b}}\left[-\exp\left(-\gamma\Big(\xi + \int_0^T (\alpha_s - a_s)\, ds\Big)\right) \,\middle|\, \mathcal{F}_t \right],$$

$$J_H(\xi_H, \alpha, a, b) := \mathbb{E}^{\mathbb{P}^{a,b}}\left[\xi_H \,\middle|\, \mathcal{F}_t \right],$$

Les fonctions de meilleure réponse de l'Agent et du Hacker s'écrivent alors respectivement :

$$v_A(\xi, \alpha, b) := \sup_{a \in \mathcal{A}(\alpha)} J_A(\alpha, a, b, \xi), \qquad v_H(\alpha, a) := \inf_{b \in \mathcal{B}} J_H(\alpha, a, b, \xi_H).$$

On utilisera ici, comme d'habitude, le concept d'équilibre de Nash pour résoudre le jeu.

Définition 1.5.30. *Pour un contrat $\xi_A \in \Xi$ offert à l'Agent, un budget $\alpha \in \mathcal{A}$, et un contrat $\xi_H \in \Xi$ offert au Hacker, un couple de stratégies $(a_0, b_0) \in \mathcal{A}(\alpha) \times \mathcal{B}$ est un* équilibre de Nash *si*

$$J_A(\xi, \alpha, a, b_0) \leq J_A(\xi, \alpha, a_0, b_0), \quad \forall a \in \mathcal{A}(\alpha),$$

$$J_H(\xi_H, \alpha, a_0, b_0) \leq J_H(\xi_H, \alpha, a_0, b), \quad \forall b \in \mathcal{B}.$$

On note $NE(\xi, \alpha, \xi_H)$ l'ensemble des équilibres de Nash associés.

Le Principal cherche à maximiser l'espérance d'un critère à travers la création d'incitations garantissant un équilibre convenable $(a^\star, b^\star)$ entre l'Agent et le Hacker. Cela conduirait au problème suivant pour le Principal :

$$\sup_{(\xi,\alpha)\in\Xi\times\mathcal{A}} \sup_{(a^\star,b^\star)\in NE(\xi,\alpha,\xi_H)} \mathbb{E}^{\mathbb{P}^{a^\star,b^\star}}\left[G\big(X_T\big) - \xi - \int_t^T c(\alpha) \right].$$

où $G : \mathbb{R}^m \to \mathbb{R}$ est une fonction de capitalisation ayant une croissance au plus exponentielle (i.e. il existe $(l, r) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ tel que $\forall x \in \mathbb{R}^m\ G(x) \leq r \exp(l\|x\|)$) et $c : \mathbb{R} \to \mathbb{R}$ est une fonction de coût strictement convexe.
Cependant, dans un contexte réaliste, ni l'Agent ni le Principal ne connaissent l'objectif exact du Hacker, et donc l'ensemble $NE(\xi, \alpha, \xi_H)$ leur est inconnu. Afin de rester en mesure de traiter le problème, on considère un scénario worst case : le Hacker est vu comme un joueur dont le but est de saboter le Principal, c'est-à-dire de minimiser sa fonction objectif. Ce point de vue revient à considérer le Hacker comme un agent travaillant pour une tierce partie, dont l'utilité consiste précisément à nuire au Principal. D'où l'appellation du jeu de Blotto Principal-Agent pessimiste.

Définition 1.5.31. *Ainsi, le problème* worst-case scenario *du Principal s'écrit :*

$$V_P := \sup_{(\xi,\alpha)\in\Xi\times\mathcal{A}} \inf_{\xi_H\in\Xi} \inf_{(a,b)\in NE(\xi,\alpha,\xi_H)} \mathbb{E}^{\mathbb{P}^{a,b}}\left[G(X_T) - \xi - \int_t^T c(\alpha) \right], \tag{1.12}$$

où $\xi_H \in \Xi$ désigne le contrat équivalent du Hacker, qui détermine la manière dont celui-ci optimise son propre critère J_H.

Résolution du jeu entre l'agent et le hacker

On passe alors, sous cette formulation pessimiste, à la résolution du jeu des agents. Pour ce faire, on utilise l'EDSR bidimensionnelle suivante :

$$Y_t^x \;=\; \begin{bmatrix}\xi_A\\ \xi_H\end{bmatrix} + \int_t^T F(Z_s,\alpha_s)\,ds - \int_t^T Z_s\,dW_s, \tag{1.13}$$

où W est un $2m$-mouvement Brownien sous P, $Z_s dW_s := (Z_s^1 dW_s^1, Z_s^2 dW_s^2)^\top$, et où le générateur $F : \mathbb{R}^{2m}\times\mathbb{R}_+ \to \mathbb{R}$ est défini par

$$F(z,\alpha) := f(z,\alpha,a_0,b_0),$$

avec

$$f : \mathbb{R}^{2m}\times\mathbb{R}_+\times\mathbb{R}^m\times\mathbb{R}^m \;\to\; \mathbb{R},$$
$$f(z,\alpha,a,b) := \begin{bmatrix} z^1\cdot(a-b) + \big(\alpha - \sum_{i=1}^m a_i\big) - \frac{\gamma}{2}\,\|z^1\|^2 \\ z^2\cdot(a-b)\end{bmatrix}.$$

Le couple (a_0,b_0) est caractérisé par

$$f(z,\alpha,a_0,b_0)_1 \;=\; \sup_{a\in\mathcal{A}(\alpha)} f(z,\alpha,a,b_0)_1, \qquad f(z,\alpha,a_0,b_0)_2 \;=\; \inf_{b\in\mathcal{B}} f(z,\alpha,a_0,b)_2.$$

Cette optimisation peut être effectuée de manière explicite, comme le montre la définition et le lemme suivants :

Définition 1.5.32. *Soit $z = (z^1,z^2)\in\mathbb{R}^{2m}$ avec $z^1,z^2\in\mathbb{R}^m$, et soit $(e_1,\dots,e_m)$ la base canonique de $\mathbb{R}^m$. On définit, pour $\alpha,\beta>0$:*

$$a^\star(z^1,\alpha) := \alpha\, e_{i^\star}\,\mathbf{1}_{\{z^1_{i^\star}\geq 1\}}, \qquad i^\star \in \arg\max_{i\in\{1,\dots,m\}} z_i^1,$$

$$b^\star(z^2,\beta) := \beta\, e_{j^\star}\,\mathbf{1}_{\{z^2_{j^\star}\geq 0\}}, \qquad j^\star \in \arg\max_{j\in\{1,\dots,m\}} z_j^2.$$

En cas d'égalité des maxima, on choisit l'indice minimal.

Lemme 1.5.33. *Avec la définition 1.5.32, le générateur F s'écrit*

$$F(z,\alpha) = f\big(z,\alpha,\, a^\star(z^1,\alpha),\, b^\star(z^2,\beta)\big),$$

La preuve de l'optimalité est explicitée dans le chapitre 3.

Remarque 1.5.34. *Les processus z^1 et z^2 représentent respectivement les sensibilités de l'Agent et du Hacker à leurs contrats. Le contrôle du Hacker consiste simplement à choisir le* champ de bataille *présentant la plus forte sensibilité (c'est-à-dire celui dont le gradient est maximal) afin d'y consacrer l'intégralité de son budget. Á l'inverse, le contrôle optimal de l'Agent consiste à travailler sur le contrat lorsque ce travail est plus rentable que le détournement du budget. Dans ce cas, l'Agent consacre lui aussi tout son budget au* champ de bataille *le plus sensible.*

Dans le même esprit de la démarche dans le cas général, on écrira la version progressive avec y de l'EDSR 1.13. Cette écriture ne permettra pas de résoudre le problème traité, contrairement au cas général, à cause de la double optimisation sur les contrats (cf Remarque 1.5.39). Elle reste néanmoins utile, notamment pour prouver le résultat principal de réduction.

Définition 1.5.35. *Pour tout $\alpha \in \mathcal{A}$, on définit l'ensemble admissible $\mathcal{Z}_{AH}^{\alpha}$ par*

$$\mathcal{Z}_{AH}^{\alpha} := \Big\{ Z \text{ prévisible à valeurs dans } \mathbb{R}^m : \ \mathbb{E}\bigg[\int_0^T \|Z_t\|^2 \mathrm{d}t\bigg] < \infty, \ \mathbb{E}\bigg[\int_0^T \big|F(Z_t, \alpha_t)\big| \mathrm{d}t\bigg] < \infty \Big\}.$$

On définit également, pour $n \in \mathbb{N}$, l'ensemble $\mathcal{Z}_{AH}^{\alpha,n} := \underbrace{\mathcal{Z}_{AH}^{\alpha} \times \cdots \times \mathcal{Z}_{AH}^{\alpha}}_{n \text{ fois}}$.

Définition 1.5.36. *Pour $y = (y^1, y^2) \in \mathbb{R}^2$, $\alpha \in \mathcal{A}$ et $Z \in \mathcal{Z}_{AH}^{\alpha,2}$, on définit le processus suivant, à valeurs dans $\mathbb{R}^2$,*

$$Y_t^{y,Z,\alpha} := y - \int_0^t F(Z_s, \alpha_s)\mathrm{d}s + \int_0^t Z_s \mathrm{d}X_s, \quad t \in [0,T]. \tag{1.14}$$

On note les coordonnées du processus ci-dessus $Y^{y,Z,\alpha,1}$ et $Y^{y,Z,\alpha,2}$.

Remarque 1.5.37. *L'ensemble $\mathcal{Z}_{AH}^{\alpha}$ est choisi de manière à garantir que les dynamiques* (1.14) *sont bien définies pour tout $Z \in \mathcal{Z}_{AH}^{\alpha,2}$, et que $Y^{y,Z,\alpha}$ est un processus adapté et de carré intégrable. En effet, les conditions imposées assurent que les deux intégrales sont bien définies, et l'EDS progressive n'est ici qu'une définition du processus.*

Si les contrats de l'agent et du hacker ont une forme commode (*i.e.* s'ils peuvent être représentés comme la valeur terminale d'un certain $Y^{y,Z,\alpha}$), alors on peut caractériser facilement les équilibres de Nash du jeu. Le résultat suivant se démontre de manière similaire à [44, Corollaire 3.31].

Proposition 1.5.38. *Soit $y \in \mathbb{R}^2$, $\alpha \in \mathcal{A}$, $Z = (Z^1, Z^2) \in \mathcal{Z}_{AH}^{\alpha,2}$ et des contrats $\xi_A = Y_T^{y,Z,\alpha,1}$, $\xi_H = Y_T^{y,Z,\alpha,2}$. On définit les processus $a_s^{\star,Z^1,\alpha} := a^\star(Z_s^1, \alpha_s)$ et $b_s^{\star,Z^2,\beta} := b^\star(Z_s^2, \beta)$, pour tout $s \in [0,T]$. Alors le couple $(a^{\star,Z^1,\alpha}, b^{\star,Z^2,\beta})$ est un équilibre de Nash du jeu entre l'agent et le hacker. De plus, la valeur de l'agent est donnée par $v_A(\xi_A, \alpha, b^{\star,Z^2,\beta}) = -e^{-\gamma y^1}$ et la valeur du hacker par $v_H(\xi_H, a^{\star,Z^1,\alpha}) = y^2$.*

Depuis les travaux en [22], on sait que, dans une interaction Principal–Agent, le problème du Principal peut être reformulé comme un problème de contrôle stochastique en restreignant, sans perte de généralité, l'ensemble des contrats aux valeurs terminales de processus contrôlés adéquats. Ce résultat a été étendu dans [27, 44] au cas de plusieurs agents, à condition que la formulation faible du problème ne fasse intervenir que des changements de mesure équivalents. Dans un tel cadre, l'étude de l'EDSR associée au problème permet de caractériser les équilibres de Nash du jeu induit et d'obtenir une paramétrisation commode des contrats admissibles. Dans notre cadre, on utilise l'approche par EDSR dans le même esprit, comme un outil de représentation : on introduit une famille tractable de contrats pour laquelle l'équilibre de Nash entre l'agent et le hacker peut être caractérisé explicitement.

Remarque 1.5.39. *Il est important de noter qu'on n'applique pas, à ce stade, les résultats de restriction des contrats issus de la littérature au couple (ξ_A, ξ_H), car dans notre modèle le hacker n'est pas engagé par le Principal mais par un tiers, et le Principal adopte une attitude pessimiste (au worst case) vis-à-vis du contrat du hacker. Appliquer directement la réduction à (ξ_A, ξ_H) nécessiterait une paramétrisation à deux volets dans laquelle les composantes correspondant à ξ_A et à ξ_H sont contrôlées par des entités différentes (autrement dit, il faudrait appliquer la réduction deux fois), ce qui conduirait à une formulation imbriquée contrôle/jeu assez lourde. On évite cela en montrant, dans la section suivante, que l'effet au worst case de ξ_H se réduit à un choix défavorable de l'action du hacker b. Après cette réduction, seul subsiste le contrat du Principal ξ_A, et on peut alors appliquer la réduction sous sa forme standard (à contrat unique).*

Élimination du contrat du Hacker

Étant donnée la solution du jeu entre l'Agent et le Hacker, le Principal peut désormais l'anticiper et reformuler son propre problème, en tenant compte des stratégies optimales que les joueurs adopteraient. Le Principal et l'Agent ne connaissent pas le contrat du Hacker ; on rappelle donc le problème wost-case du Principal :

$$V_P = \sup_{(\xi_A,\alpha)\in\Xi\times\mathcal{A}} \inf_{\xi_H\in\Xi} \inf_{(a,b)\in NE(\xi_A,\alpha,\xi_H)} \mathbb{E}^{P^{a,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\mathrm{d}s\right].$$

Compte tenu de la formulation du problème, on peut observer que ξ_H n'est utilisé que pour déterminer le contrôle b du Hacker. Ainsi, pour un espace de contrats Ξ suffisamment riche, on peut reformuler le problème et optimiser directement sur b, comme on le montre dans cette section : c'est ce qu'on appellera *l'élimination du contrat du hacker.*

On introduit d'abord :

$$J_P(\xi_A,\alpha;a,b) := \mathbb{E}^{P^{a,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\mathrm{d}s\right].$$

On a besoin, pour cette section, d'écrire les équilibres de Nash sous une forme légèrement différente, tout en conservant la même définition.

Definition 1.5.40. *Pour $(\xi_A,\alpha)\in\Xi\times\mathcal{A}$ fixé, on définit la correspondance de meilleure réponse de l'Agent à un contrôle b du Hacker :*

$$BR_A(\xi_A,\alpha;b) := \arg\max_{a\in\mathcal{A}(\alpha)} J_A(\xi_A,\alpha;a,b),$$

et, pour $\xi_H\in\Xi$ fixé, la correspondance de meilleure réponse du Hacker à un contrôle a de l'Agent :

$$BR_H(\xi_H;a) := \arg\min_{b\in B} J_H(\xi_H;a,b).$$

Un couple $(a,b)\in\mathcal{A}(\alpha)\times\mathcal{B}$ est un équilibre de Nash du jeu Agent–Hacker si

$$a\in BR_A(\xi_A,\alpha;b) \qquad et \qquad b\in BR_H(\xi_H;a).$$

L'ensemble des équilibres de Nash $NE(\xi_A,\alpha,\xi_H)$ n'est pas modifié par cette caractérisation.

Hypothèse 1.5.41 (Implémentabilité d'une action worst-case du hacker)**.** *Pour tout (ξ_A,α), il existe $\xi_H^\dagger\in\Xi$ et $(a^\dagger,b^\dagger)\in NE(\xi_A,\alpha,\xi_H^\dagger)$ tels que*

$$J_P(\xi_A,\alpha;a^\dagger,b^\dagger) = \inf_{(a,b)\in NE(\xi_A,\alpha,\xi_H^\dagger)} J_P(\xi_A,\alpha;a,b) \quad et \quad b^\dagger\in\arg\min_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A,\alpha;a,b).$$

L'hypothèse 1.5.41 signifie qu'on peut implémenter (via un certain contrat $\xi_H^\dagger$) une action du hacker $b^\dagger$ qui atteint la valeur worst-case pour le Principal.

Proposition 1.5.42. *On fixe $(\xi_A,\alpha)\in\Xi\times\mathcal{A}$ et on suppose que le problème de minimisation interne après élimination du contrat du Hacker admet un minimiseur $b^\dagger\in\mathcal{B}_{\text{ext}}$, c'est-à-dire*

$$b^\dagger\in\arg\min_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A,\alpha;a,b), \qquad B_{\text{ext}} = \{0,\beta e_1,\dots,\beta e_m\}.$$

On définit le processus borné, $\mathbb{F}$-progressivement mesurable, $Z^{2,\dagger}$ par

$$Z_t^{2,\dagger} := \begin{cases} e_k, & \text{si } b_t^\dagger = \beta e_k \text{ pour un certain } k \in \{1,\dots,m\}, \\ -\mathbf{1}, & \text{si } b_t^\dagger = 0, \end{cases} \qquad 0 \le t \le T.$$

On suppose en outre qu'il existe $y^1 \in \mathbb{R}$ et $Z^{1,\dagger} \in \mathcal{Z}^\alpha$ tels que, en posant $Z^\dagger := (Z^{1,\dagger}, Z^{2,\dagger})$ et $y := (y^1, y^2)$,

$$\xi_A = Y_T^{y,Z^\dagger,\alpha,1} \qquad \text{pour un certain } y^2 \in \mathbb{R}.$$

On définit le contrat du hacker par

$$\xi_H^\dagger := Y_T^{y,Z^\dagger,\alpha,2}.$$

Alors :

1. *$b^\star(Z_t^{2,\dagger}, \beta) = b_t^\dagger$ pour p.s. pour tout t, donc $b^\dagger$ est le contrôle d'équilibre du hacker induit par $\xi_H^\dagger$;*
2. *pour tout choix de $a^\dagger \in BR_A(\xi_A, \alpha; b^\dagger)$, le couple $(a^\dagger, b^\dagger)$ appartient à $NE(\xi_A, \alpha, \xi_H^\dagger)$;*
3. *$J_P(\xi_A, \alpha; a^\dagger, b^\dagger) = \inf_{(a,b)\in NE(\xi_A,\alpha,\xi_H^\dagger)} J_P(\xi_A, \alpha; a, b)$.*

La preuve est explicitée dans le Chapitre 3.

Cette proposition donne immédiatement le corollaire suivant :

Corollaire 1.5.43 (Vérification de l'hypothèse 1.5.41)**.** *On fixe $(\xi_A, \alpha) \in \Xi \times \mathcal{A}$. On suppose que le problème de minimisation interne après élimination du contrat du hacker admet un minimiseur $b^\dagger \in \mathcal{B}_{\text{ext}}$ et que la condition de représentation feedback de la Proposition 1.5.42 est satisfaite pour (ξ_A, α). Alors l'hypothèse 1.5.41 est vérifiée pour (ξ_A, α). En particulier, si ces hypothèses sont satisfaites pour tout (ξ_A, α), alors l'hypothèse 1.5.41 est vérifiée.*

Problème interne après élimination.

Définition 1.5.44. *Pour (ξ_A, α) fixé, on définit la valeur interne au worst case du Principal par*

$$\mathcal{W}(\xi_A, \alpha) := \inf_{\xi_H \in \Xi} \inf_{(a,b)\in NE(\xi_A,\alpha,\xi_H)} J_P(\xi_A, \alpha; a, b). \tag{1.15}$$

On définit la valeur après élimination, lorsque le contrôleur adverse choisit directement b dans la classe admissible complète $\mathcal{B}$:

$$\widetilde{\mathcal{W}}(\xi_A, \alpha) := \inf_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A, \alpha; a, b). \tag{1.16}$$

Enfin, on définit la valeur après élimination restreinte

$$\widetilde{\mathcal{W}}_{\text{ext}}(\xi_A, \alpha) := \inf_{b\in\mathcal{B}_{\text{ext}}} \inf_{a\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A, \alpha; a, b), \tag{1.17}$$

où $\mathcal{B}_{\text{ext}}$ désigne la classe des processus $\mathbb{F}$-progressivement mesurables à valeurs dans B_{ext}.

Proposition 1.5.45 (Élimination du contrat du hacker)**.** *Pour tout* $(\xi_A, \alpha) \in \Xi \times \mathcal{A}$*, si l'hypothèse 1.5.41 est vérifiée, alors*

$$\mathcal{W}(\xi_A, \alpha) = \widetilde{\mathcal{W}}(\xi_A, \alpha).$$

Par conséquent, le problème robuste du Principal

$$V_P := \sup_{(\xi_A,\alpha)\in\Xi\times\mathcal{A}} \mathcal{W}(\xi_A, \alpha)$$

admet la formulation équivalente

$$V_P = \sup_{(\xi_A,\alpha)\in\Xi\times\mathcal{A}} \inf_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} \mathbb{E}^{\mathbb{P}^{a,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\, ds\right].$$

La preuve est explicitée dans le Chapitre 3.

Ce résultat est fondamental pour l'analyse des contributions du manuscrit. L'idée de la Proposition 1.5.45 est de montrer que, sous l'Hypothèse 1.5.41, on peut réduire le problème d'optimisation initial — posé sur un espace de contrats — à un problème d'optimisation portant directement sur les contrôles du Hacker. De son côté, le Corollaire 1.5.43 assure que l'Hypothèse 1.5.41 est satisfaite pour une classe particulière de modèles. On montrera dans la partie suivante que le problème de Principal–Agent de Blotto appartient à cette classe, ce qui permettra ensuite de le résoudre via la formulation réduite, sans avoir à considérer des optimisations imbriquées sur les contrats.

Résolution du problème du principal

Dans la section précédente, on a introduit le critère interne après élimination

$$\widetilde{\mathcal{W}}(\xi_A, \alpha) = \inf_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A, \alpha; a, b),$$

et on a montré (sous l'Hypothèse 1.5.41) que, pour tout (ξ_A, α),

$$\mathcal{W}(\xi_A, \alpha) = \widetilde{\mathcal{W}}(\xi_A, \alpha).$$

Dans cette section, on *part d'emblée du problème du Principal après élimination* et on en dérive la caractérisation HJB. On note sa valeur

$$\widetilde{V}_P := \sup_{(\xi_A,\alpha)\in\Xi\times\mathcal{A}} \widetilde{\mathcal{W}}(\xi_A, \alpha) = \sup_{(\xi_A,\alpha)\in\Xi\times\mathcal{A}} \inf_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} \mathbb{E}^{\mathbb{P}^{a,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\, ds\right]. \tag{1.18}$$

À la fin de la section, on identifie un contrôle explicite de worst case $b^\dagger$ dans le Hamiltonien et on montre qu'on peut le choisir à valeurs dans B_{ext}. Cela permet d'appliquer la Proposition 1.5.42 et le Corollaire 1.5.43 pour vérifier l'Hypothèse 1.5.41 dans notre cadre. En combinant cette vérification avec la Proposition 1.5.45, on conclut que la valeur après élimination $\widetilde{V}_P$ coïncide avec la valeur originale du Principal V_P.

Puisque le contrat du Hacker a été éliminé, on applique la machinerie de restriction des contrats sous sa forme standard *à contrat unique*, c'est-à-dire uniquement au contrat Principal–Agent ξ_A. On obtient ainsi une formulation en contrôle stochastique en termes de l'utilité de réserve de l'agent, ce qui conduit à une équation HJBI.

1.5.3 Dérivation de l'EDP HJBI

On dérive maintenant l'équation HJBI caractérisant la valeur après élimination $\widetilde{V}_P$ définie en (1.18). Dans toute la suite, on se place sous la formulation faible et on considère la valeur du problème.

Définition 1.5.46. *Pour $(z, \alpha) \in \mathbb{R}^m \times \mathbb{R}_+$, on définit*

$$F_0(z, \alpha) := \sup_{a \in \mathcal{A}(\alpha)} \Big\{ z \cdot a + \alpha - \mathbf{1}^\top a \Big\} - \frac{\gamma}{2} \|z\|^2, \tag{1.19}$$

et on définit la meilleure réponse explicite correspondante

$$a^*(z, \alpha) := \alpha \, e_i \, \mathbb{1}_{\{z_i \geq 1\}}, \qquad \text{où } i \in \arg\max_{j=1,\dots,m} z \cdot e_j. \tag{1.20}$$

De manière équivalente,

$$F_0(z, \alpha) = z \cdot a^*(z, \alpha) + \alpha - \mathbf{1}^\top a^*(z, \alpha) - \frac{\gamma}{2} \|z\|^2. \tag{1.21}$$

Remarque 1.5.47. *Pour (z, α, b) fixés, le terme $-z \cdot b$ est constant en a, d'où*

$$\arg\max_{a \in \mathcal{A}(\alpha)} \{ z \cdot (a - b) + \alpha - \mathbf{1}^\top a \} = \arg\max_{a \in \mathcal{A}(\alpha)} \{ z \cdot a + \alpha - \mathbf{1}^\top a \}.$$

En particulier, le contrôle feedback $a^*(z, \alpha)$ défini en (1.20) *appartient à $BR_A(\xi_A, \alpha; b)$.*

On introduit maintenant la classe d'admissibilité usuelle garantissant que le processus de contrat feedback est bien défini.

Définition 1.5.48. *Pour tout $\alpha \in \mathcal{A}$, on définit*

$$\begin{aligned} \mathcal{Z}^\alpha := \Big\{ Z \ \mathbb{F}\text{-progressivement mesurable, à valeurs dans } \mathbb{R}^m \ : \mathbb{E}\Big[\int_0^T \|Z_t\|^2 \, dt \Big] < \infty, \\ \mathbb{E}\Big[\int_0^T \big| F_0(Z_t, \alpha_t) \big| \, dt \Big] < \infty \Big\}. \end{aligned} \tag{1.22}$$

Remarque 1.5.49. *Bien que b n'apparaisse pas dans la définition de F_0, il intervient dans les fonctionnelles de performance via $\int_0^T Z_t \cdot b_t \, dt$. Ce terme est automatiquement intégrable uniformément en $b \in \mathcal{B}$ dès que $Z \in L^2(dt \otimes d\mathbb{P})$. En particulier, dans le cadre Blotto, $\|b_t\| \leq \beta$ pour tout b admissible, et pour tout $Z \in \mathcal{Z}^\alpha$ et tout $b \in \mathcal{B}$,*

$$\mathbb{E}\Big[\int_0^T |Z_t \cdot b_t| \, dt \Big] \leq \beta \sqrt{T} \Big(\mathbb{E} \int_0^T \|Z_t\|^2 dt \Big)^{1/2} < \infty.$$

Ainsi, on n'a pas besoin d'incorporer b dans la définition de $\mathcal{Z}^\alpha$.

Définition 1.5.50 (Contrats Feedback après élimination)**.** *Pour $y \in \mathbb{R}$, $\alpha \in \mathcal{A}$ et $Z \in \mathcal{Z}^\alpha$, on définit le processus unidimensionnel*

$$Y_t^{y,Z,\alpha} := y - \int_0^t F_0(Z_s, \alpha_s) \, ds + \int_0^t Z_s \, dX_s, \qquad t \in [0, T], \tag{1.23}$$

et on pose le contrat Principal–Agent $\xi_A := Y_T^{y,Z,\alpha}$.

Remarque 1.5.51. *Dans le cadre en formulation faible de [22], on peut, sans perte de généralité, se restreindre à des contrats Principal–Agent ξ_A admettant une représentation feedback de la forme* (1.23)*. En particulier, lors de la vérification de l'Hypothèse 1.5.41 ci-dessous, il suffit de considérer de tels contrats feedback. De plus, dès qu'un contrôle de worst case borné $b^\dagger \in \mathcal{B}_{\text{ext}}$ a été identifié, cette restriction fournit la "représentabilité en Z^1" requise dans la Proposition 1.5.42 : en effet, pour α fixé et $b^\dagger$ borné, on peut choisir $Z^{1,\dagger}$ de sorte que ξ_A soit la valeur terminale de la première composante dans la représentation bidimensionnelle associée à $(Z^{1,\dagger}, Z^{2,\dagger})$.*

On note R_0 l'utilité de réserve de l'agent. Dans notre cadre, la valeur de l'agent sous le contrat feedback $\xi_A = Y_T^{y,Z,\alpha}$ ne dépend de (y, Z, α) qu'au travers de y (comme dans la Proposition 1.5.38, désormais à une seule composante). En particulier, la contrainte de participation est équivalente à $y \geq \bar{y}(R_0)$ et est donc saturée à l'optimum. Avec un léger abus de notation, on note encore R_0 le niveau de réserve, de sorte que la condition saturée s'écrive $y = R_0$.

Valeur du Principal après élimination en forme feedback. En utilisant la paramétrisation feedback ci-dessus et la meilleure réponse $a^*(Z, \alpha)$, on peut réécrire (1.18) sous la forme

$$\widetilde{V}_P := \sup_{\alpha \in \mathcal{A}} \sup_{Z \in \mathcal{Z}^\alpha} \inf_{b \in \mathcal{B}} \mathbb{E}^{\mathbb{P}^{a^*(Z,\alpha),b}} \left[G(X_T) - Y_T^{R_0,Z,\alpha} - \int_0^T c(\alpha_s)\, ds \right]. \tag{1.24}$$

Définition 1.5.52 (Fonction valeur)**.** *Pour $(t, x, y) \in [0, T] \times \mathbb{R}^m \times \mathbb{R}$, on définit la fonction valeur du Principal après élimination par*

$$v(t, x, y) := \sup_{\alpha \in \mathcal{A}} \sup_{Z \in \mathcal{Z}^\alpha} \inf_{b \in \mathcal{B}} \mathbb{E}^{\mathbb{P}^{a^*(Z,\alpha),b}} \left[G(X_T) - Y_T^{t,x,y;Z,\alpha} - \int_t^T c(\alpha_s)\, ds \right], \tag{1.25}$$

où, pour des contrôles donnés (α, Z, b), le couple $(X^{t,x}, Y^{t,x,y;Z,\alpha})$ est défini comme suit :

- *$X^{t,x}$ part de $X_t^{t,x} = x$ et évolue sous $\mathbb{P}^{a^*(Z,\alpha),b}$ selon la formulation faible, i.e. $W^{a,b} := X - \int_0^\cdot (a^*(Z_s, \alpha_s) - b_s) ds$ est un mouvement brownien et, par conséquent,*

$$dX_s^{t,x} = (a^*(Z_s, \alpha_s) - b_s)\, ds + dW_s^{a,b}, \qquad s \in [t, T];$$

- *$Y^{t,x,y;Z,\alpha}$ part de $Y_t^{t,x,y;Z,\alpha} = y$ et est donné par*

$$Y_s^{t,x,y;Z,\alpha} := y - \int_t^s F_0(Z_r, \alpha_r)\, dr + \int_t^s Z_r \cdot dX_r^{t,x}, \qquad s \in [t, T],$$

de sorte que le contrat terminal est $\xi_A = Y_T^{t,x,y;Z,\alpha}$.

Comme $Y_T^{y+\delta,Z,\alpha} = Y_T^{y,Z,\alpha} + \delta$ pour tout $\delta \in \mathbb{R}$, on a la propriété de translation $v(t, x, y+\delta) = v(t, x, y) - \delta$. Il existe donc une fonction $u : [0, T] \times \mathbb{R}^m \to \mathbb{R}$ telle que

$$v(t, x, y) = u(t, x) - y, \qquad \text{et en particulier} \qquad \widetilde{V}_P(t, x) = v(t, x, R_0) = u(t, x) - R_0. \tag{1.26}$$

Sous les hypothèses standards assurant le principe de programmation dynamique pour le jeu à somme nulle après élimination et la réduction à contrat unique de [22, 44], u est caractérisée comme solution de viscosité de l'EDP ci-dessous.

Équation HJB. Dans le cas markovien, la programmation dynamique appliquée formellement à (1.24) donne une équation HJB pour u :

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) + \sup_{\alpha\in A}\sup_{z\in\mathbb{R}^m}\inf_{b\in B}\Big\{(a^*(z,\alpha)-b)\cdot(\nabla u(t,x)-z) + F_0(z,\alpha) - c(\alpha)\Big\} = 0,$$
$$u(T,x) = G(x). \tag{1.27}$$

En utilisant (1.21), on réécrit le terme entre accolades sous la forme

$$a^*(z,\alpha)\cdot\nabla u(t,x) - b\cdot(\nabla u(t,x)-z) + \alpha - \mathbf{1}^\top a^*(z,\alpha) - \frac{\gamma}{2}\|z\|^2 - c(\alpha).$$

Remarque 1.5.53. *Il convient de souligner que l'équation de Hamilton–Jacobi–Bellman ci-dessus est ici utilisée à un niveau formel. En effet, une justification rigoureuse par programmation dynamique est délicate dans le cadre étudié, en raison notamment de la non-compacité des espaces de contrôle. Des travaux récents, en particulier ceux de Hernández-Santibáñez et Mastrolia [43], montrent toutefois qu'une telle caractérisation peut souvent être retrouvée par une approche de type Perron stochastique couplée à un principe de comparaison pour l'EDP associée. On pense, sans le démontrer rigoureusement dans ce manuscrit, qu'une telle stratégie pourrait également être adaptée au jeu considéré ici, compte tenu de la structure explicite du Hamiltonien réduit.*

Vérification de l'élimination du contrat du hacker.

Définition 1.5.54. *On fixe $(\xi_A,\alpha)\in\Xi\times\mathcal{A}$. On suppose que ξ_A est donné sous forme feedback par la Définition 1.5.50 :*

$$\xi_A \;=\; Y_T^{R_0,Z^A,\alpha},$$

pour un certain $Z^A\in\mathcal{Z}^\alpha$, et on définit la meilleure réponse associée de l'agent

$$a_t^* := a^*(Z_t^A,\alpha_t), \qquad 0\le t\le T.$$

On définit le processus adapté

$$\kappa_t \;:=\; F_0(Z_t^A,\alpha_t) - c(\alpha_t) \;=\; Z_t^A\cdot a_t^* + \alpha_t - \mathbf{1}^\top a_t^* - \frac{\gamma}{2}\|Z_t^A\|^2 - c(\alpha_t), \qquad 0\le t\le T.$$

Pour tout contrôle admissible du hacker $b\in\mathcal{B}$, on définit le générateur (Lipschitzien)

$$f^b(t,z) \;:=\; (a_t^*-b_t)\cdot(z-Z_t^A) \;+\; \kappa_t, \qquad (t,z)\in[0,T]\times\mathbb{R}^m,$$

et l'EDSR correspondante sous la mesure de référence $\mathbb{P}$ (de sorte que X est le mouvement brownien canonique)

$$Y_t^b = G(X_T) + \int_t^T f^b(s,Z_s^b)\,ds - \int_t^T Z_s^b\cdot dX_s, \qquad 0\le t\le T. \tag{1.28}$$

On définit aussi le générateur enveloppe

$$f(t,z) \;:=\; \inf_{b\in B}\big\{(a_t^*-b)\cdot(z-Z_t^A)\big\} + \kappa_t \;=\; a_t^*\cdot(z-Z_t^A) \;-\; \beta\max\Big(0,\max_{1\le i\le m}(z_i-(Z_t^A)_i)\Big) \;+\; \kappa_t,$$

et l'EDSR enveloppe

$$Y_t = G(X_T) + \int_t^T f(s,Z_s)\,ds - \int_t^T Z_s\cdot dX_s, \qquad 0\le t\le T. \tag{1.29}$$

Définition 1.5.55. *Dans le cadre de la Définition 1.5.54, soit* (Y, Z) *une solution de* (1.29)*. On définit le processus* B_{ext}*-valué, progressivement mesurable,* $b^\dagger$ *par*

$$b_t^\dagger := \begin{cases} 0, & \text{si } \max_{1\le i\le m}\left(Z_{t,i} - (Z_t^A)_i\right) \le 0, \\ \beta e_{i^*(t)}, & \text{sinon, où } i^*(t) = \min\arg\max_{1\le i\le m}\left(Z_{t,i} - (Z_t^A)_i\right). \end{cases}$$

Proposition 1.5.56 (Construction d'un contrôle de pire cas à valeurs dans B_{ext})**.** *On se place sous les hypothèses de la Définition 1.5.54. Alors :*

1. *Pour tout* $b \in \mathcal{B}$*, l'EDSR* (1.28) *admet une unique solution* (Y^b, Z^b)*, et l'EDSR enveloppe* (1.29) *admet une unique solution* (Y, Z)*.*
2. *Pour tout* $b \in \mathcal{B}$*, la valeur initiale vérifie*

$$Y_0^b - R_0 \;=\; \mathbb{E}^{\mathbb{P}^{a^*,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\,ds\right] \;=\; J_P(\xi_A, \alpha; a^*, b).$$

3. *Le processus* $b^\dagger$ *défini dans la Définition 1.5.55 appartient à* $\mathcal{B}_{\text{ext}}$ *et est un minimiseur global du problème interne après élimination pour le couple fixé* (ξ_A, α) *:*

$$J_P(\xi_A, \alpha; a^*, b^\dagger) \;=\; \inf_{b\in\mathcal{B}} J_P(\xi_A, \alpha; a^*, b) \;=\; \inf_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A, \alpha; a, b).$$

La preuve s'appuie sur un argument de comparaison d'EDSR et est explicitée dans le Chapitre 3.

Proposition 1.5.57. *Dans le jeu Principal–Agent de Blotto, les valeurs après élimination et originale coïncident :*

$$\widetilde{V}_P = V_P.$$

La preuve, explicitée dans le Chapitre 3, s'appuie sur le fait que la Proposition 1.5.56 fournit un contrôle de worst case $b^\dagger$ sous forme feedback. Avec la restriction feedback sur ξ_A (Remarque 1.5.51), la Proposition 1.5.42 et le Corollaire 1.5.43 permettent de vérifier l'Hypothèse 1.5.41. En particulier, en combinant avec la Proposition 1.5.45, on obtient le résultat.

Optimisation du Hamiltonien

Définition 1.5.58. *On définit pour tout* $t \ge 0$

$$\mathcal{R}_k := \{z \in \mathbb{R}^m :\; z_k = \max_j z_j,\; z_k \ge 1\},$$

et

$$z_i^{(k)}(t) := (p_i - t)_+, \qquad i \ne k, \qquad z_k^{(k)}(t) := \max\left\{1,\; \max_{1\le j\le m}(p_j - t)_+\right\} = \max\{1, p_{\max} - t\}. \tag{1.30}$$

Après l'optimisation sur z, en notant $p := \nabla u(t,x)$, le Hamiltonien s'écrit de la manière suivante

$$H(p) = \sup_{\alpha\ge 0} \max\left\{H_0(p;\alpha),\; \max_{1\le k\le m} H_k(p;\alpha)\right\}, \tag{1.31}$$

où

$$H_0(p;\alpha) := \alpha - c(\alpha) \; - \inf_{t \geq \max(0, p_{\max}-1)} \left\{ \frac{\gamma}{2} \sum_{i=1}^{m} (p_i - t)_+^2 + \beta t \right\}, \tag{1.32}$$

$$H_k(p;\alpha) := \alpha p_k - c(\alpha) - \inf_{t \geq 0} \left\{ \frac{\gamma}{2} \Big((z_k^{(k)}(t))^2 + \sum_{i \neq k} (p_i - t)_+^2 \Big) + \beta t \right\}, \tag{1.33}$$

où $z_k^{(k)}(t)$ correspond à la définition donnée en (1.30). Les valeurs optimales sont explicitement trouvées dans le Chapitre 3

La représentation 1.31 met en évidence une décomposition en deux régimes. Le terme $H_0(p;\alpha)$ correspond à un régime *sans investissement* : les incitations demeurent en dessous du seuil d'activation, de sorte que l'agent n'est pas incité à concentrer son budget sur un unique champ de bataille, et l'effet adverse se réduit à la pénalité $\beta\, t$. Le scalaire t admet une interprétation naturelle comme un *niveau de plafonnement* pour la quantité $p - z$: minimiser en t revient à choisir le sabotage minimal compatible avec un coût quadratique $\|z\|^2$, ce qui conduit à la structure de type *water-filling* $z_i = (p_i - t)_+$.

À l'inverse, les termes $H_k(p;\alpha)$ décrivent un régime *avec investissement* : les incitations sont suffisamment fortes pour déclencher une allocation concentrée sur le champ de bataille k (via la contrainte $z_k \geq 1$), et le Principal compare, pour chaque k, le gain marginal αp_k au coût instantané $c(\alpha)$, tout en payant à la fois le terme βt et la pénalité quadratique en z. Après optimisation en z, l'optimisation ponctuelle restante est unidimensionnelle en α.

Ainsi, après l'optimisation en z, l'optimisation ponctuelle restante est unidimensionnelle en α, et sa forme explicite dépend du choix du coût instantané c. Dans la suite, on donne deux exemples standards.

Coût instantané quadratique

On résout d'abord le cas d'une fonction de coût quadratique

$$c(\alpha) = \frac{\lambda}{2}\alpha^2, \qquad \lambda > 0.$$

Pour $\theta \in \mathbb{R}$, on définit le problème de maximisation à pente unique

$$\Psi_{\text{quad}}(\theta) := \sup_{\alpha \geq 0} \Big\{ \alpha\, \theta - \frac{\lambda}{2}\alpha^2 \Big\}. \tag{1.34}$$

Un calcul direct donne

$$\alpha_{\text{quad}}(\theta) = \frac{\theta_+}{\lambda}, \qquad \Psi_{\text{quad}}(\theta) = \frac{\theta_+^2}{2\lambda}, \qquad \theta_+ := \max(\theta, 0). \tag{1.35}$$

On applique maintenant ceci à (1.31) :

— Dans la branche H_0, la pente vaut $\theta = 1$, donc $\sup_{\alpha \geq 0}\{\alpha - c(\alpha)\} = \Psi_{\text{quad}}(1) = \frac{1}{2\lambda}$. On obtient donc

$$\sup_{\alpha \geq 0} H_0(p;\alpha) = \frac{1}{2\lambda} - \inf_{t \geq \max(0, p_{\max}-1)} \left\{ \frac{\gamma}{2} \sum_{i=1}^{m} (p_i - t)_+^2 + \beta t \right\}. \tag{1.36}$$

— Dans la branche H_k, la pente vaut $\theta = p_k$. Ainsi, maximiser en α donne

$$\Psi_{\text{quad}}(p_k) = \frac{(p_k)_+^2}{2\lambda},$$

et l'on obtient

$$\sup_{\alpha\geq 0} H_k(p;\alpha) = \frac{(p_k)_+^2}{2\lambda} - \inf_{t\geq 0}\left\{\frac{\gamma}{2}\Big((z_k^{(k)}(t))^2 + \sum_{i\neq k}(p_i - t)_+^2\Big) + \beta t\right\}. \tag{1.37}$$

Finalement, le Hamiltonien associé au coût quadratique est

$$H_{\text{quad}}(p) = \max\Big\{\sup_{\alpha\geq 0} H_0(p;\alpha),\ \max_{1\leq k\leq m}\sup_{\alpha\geq 0} H_k(p;\alpha)\Big\},$$

avec (1.36) et (1.37).

Coût instantané exponentiel

On résout maintenant le cas d'un coût de forme exponentielle

$$c(\alpha) = \lambda_1 e^{\lambda_2\alpha}, \qquad \lambda_1 > 0,\ \lambda_2 > 0.$$

Pour $\theta \in \mathbb{R}$, on définit le problème de maximisation à pente unique

$$\Psi_{\text{exp}}(\theta) := \sup_{\alpha\geq 0}\Big\{\alpha\,\theta - \lambda_1 e^{\lambda_2\alpha}\Big\}. \tag{1.38}$$

Comme l'objectif est strictement concave en α, le maximiseur est caractérisé par la condition du premier ordre

$$\theta - \lambda_1\lambda_2 e^{\lambda_2\alpha} = 0,$$

lorsque le maximiseur est intérieur. On en déduit l'optimiseur et la valeur explicites :

$$\alpha_{\text{exp}}(\theta) = \begin{cases} 0, & \theta \leq \lambda_1\lambda_2, \\ \frac{1}{\lambda_2}\ln\Big(\frac{\theta}{\lambda_1\lambda_2}\Big), & \theta > \lambda_1\lambda_2, \end{cases} \qquad \Psi_{\text{exp}}(\theta) = \begin{cases} -\lambda_1, & \theta \leq \lambda_1\lambda_2, \\ \frac{\theta}{\lambda_2}\Big(\ln\frac{\theta}{\lambda_1\lambda_2} - 1\Big), & \theta > \lambda_1\lambda_2. \end{cases} \tag{1.39}$$

On applique maintenant ceci à (1.31) :

— Dans la branche H_0, la pente vaut $\theta = 1$, donc $\sup_{\alpha\geq 0}\{\alpha - c(\alpha)\} = \Psi_{\text{exp}}(1)$ et

$$\sup_{\alpha\geq 0} H_0(p;\alpha) = \Psi_{\text{exp}}(1) - \inf_{t\geq\max(0,p_{\max}-1)}\left\{\frac{\gamma}{2}\sum_{i=1}^m (p_i - t)_+^2 + \beta t\right\}. \tag{1.40}$$

— Dans la branche H_k, la pente vaut $\theta = p_k$, d'où

$$\sup_{\alpha\geq 0} H_k(p;\alpha) = \Psi_{\text{exp}}(p_k) - \inf_{t\geq 0}\left\{\frac{\gamma}{2}\Big((z_k^{(k)}(t))^2 + \sum_{i\neq k}(p_i - t)_+^2\Big) + \beta t\right\}. \tag{1.41}$$

Finalement,

$$H_{\text{exp}}(p) = \max\Big\{\sup_{\alpha\geq 0} H_0(p;\alpha),\ \max_{1\leq k\leq m}\sup_{\alpha\geq 0} H_k(p;\alpha)\Big\},$$

avec (1.40) et (1.41).

EDP résultante.

Dans les deux cas, la valeur markovienne du Principal u résout l'équation réduite

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) + H(\nabla u(t,x)) = 0, \qquad u(T,x) = G(x),$$

avec $H = H_{\text{quad}}$ ou $H = H_{\exp}$.

Représentation par EDSR.

Soit $(W_t)_{0\le t\le T}$ un mouvement brownien $\mathbb{R}^m$-valué, et on définit le processus progressif

$$X_t := x_0 + W_t, \qquad 0 \le t \le T.$$

Soit (Y, Z) la solution de l'EDSR

$$Y_t = G(X_T) + \int_t^T H(Z_s)\, ds - \int_t^T Z_s \cdot dW_s, \qquad 0 \le t \le T. \tag{1.42}$$

Alors $u(0, x_0) = Y_0$. Plus généralement, en posant $u(t,x) := Y_t^{t,x}$, on obtient une solution de viscosité de

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) + H(\nabla u(t,x)) = 0, \qquad u(T,x) = G(x).$$

Cette EDSR est résolue numériquement dans la section suivante.

Résultats numériques

On illustre maintenant l'évolution temporelle du problème Principal–Agent après élimination sur des trajectoires représentatives. On utilise pour ces résultats le schéma Deep BSDE modifié présenté dans la section numérique 1.4.9. Dans toute cette section, on fixe $m = 10$, $x_0 = 0_{\mathbb{R}^{10}}$ et $T = 1$. On prend un coût instantané quadratique

$$c(\alpha) = \frac{1}{4}\alpha^2,$$

et un gain terminal de type Blotto

$$G(x) = S \sum_{i=1}^{m} w_i \, \mathbf{1}_{\{x_i \ge 0\}}, \qquad w_1 = 2, \;\; w_i = 1 \;\; (i = 2, \dots, m),$$

où $S > 0$ est un paramètre d'échelle qui représente les enjeux. On résout l'EDSR associée à l'EDP semi-linéaire

$$\partial_t u(t,x) + \tfrac{1}{2}\Delta u(t,x) + H(\nabla u(t,x)) = 0, \qquad u(T,x) = G(x),$$

On utilise un réseau pour Z avec deux couches cachées de taille 500 et une activation SiLU. On effectue un warm-start avec un réseau initialisé aléatoirement et $y_0 = S \times \frac{\sum_i w_i}{2}$. Après le warm-start, on applique une relaxation du générateur avec $\delta_i = \frac{i}{5}$, $i = 1, 2, 3, 4, 5$. Pour chaque valeur δ_i, on entraîne en trois phases :

1. une première phase de 800 itérations avec $N = 120$ pas de temps, $\lambda_{\text{res}} = 0.02$ et $\lambda_Z = 0.05$;
2. une deuxième phase de 800 itérations avec $N = 160$ pas de temps, $\lambda_{\text{res}} = 0.015$ et $\lambda_Z = 0.03$;

3. une phase finale de 1000 itérations avec $N = 200$ pas de temps, $\lambda_{\text{res}} = 0.01$ et $\lambda_Z = 0.02$.

Le long d'une trajectoire brownienne simulée $X_t = x_0 + W_t$, le solveur Deep-BSDE renvoie des processus $(Y_t, Z_t)_{0 \le t \le T}$, où $Y_t$approxime la valeur markovienne $u(t, X_t)$, et où l'on a $Z_t \approx \nabla u(t, X_t)$. À chaque instant t, on reconstruit le budget instantané optimal du Principal en maximisant le Hamiltonien d'Isaacs au gradient courant $p_t := Z_t$, i.e.

$$\alpha_t^\star = \alpha^\star(p_t),$$

en utilisant les règles en forme fermée dérivées dans la Section 1.5.3 et la décomposition en branches $H = \max\{H_0, H_1\}$.

Remarque 1.5.59. *Une implication qualitative clé de la décomposition du Hamiltonien est l'existence de deux régimes. Sous un coût quadratique, dès que la branche sans investissement H_0 domine, le maximiseur ponctuel en α est constant : $\alpha_t^\star = 1/\lambda = 2$. Ainsi, un plateau de $\alpha_t^\star$ constitue une signature structurelle du régime H_0. Lorsque le paramètre d'enjeux S augmente, les gradients deviennent plus grands et la branche d'investissement H_1 devient optimale plus fréquemment, ce qui produit des budgets dépendants de l'état et variables dans le temps.*

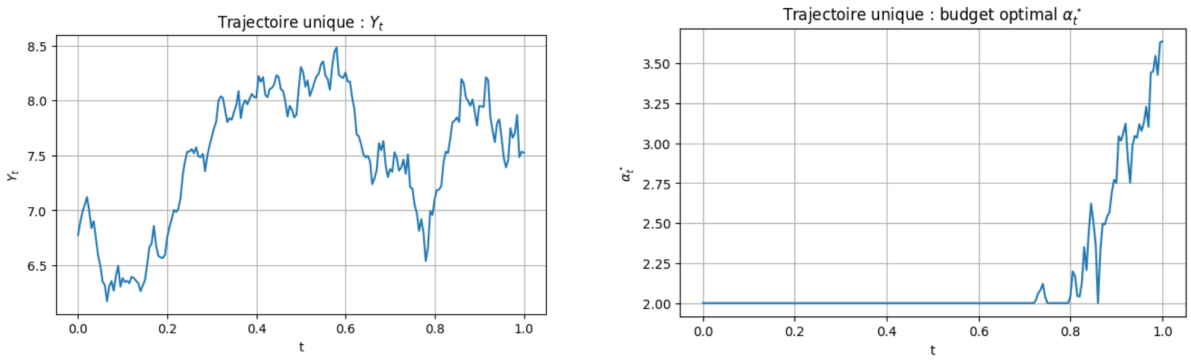


FIGURE 1.6 – Régime à faibles enjeux ($S = 1$) : trajectoires de Y_t et budget optimal reconstruit $\alpha_t^\star$.

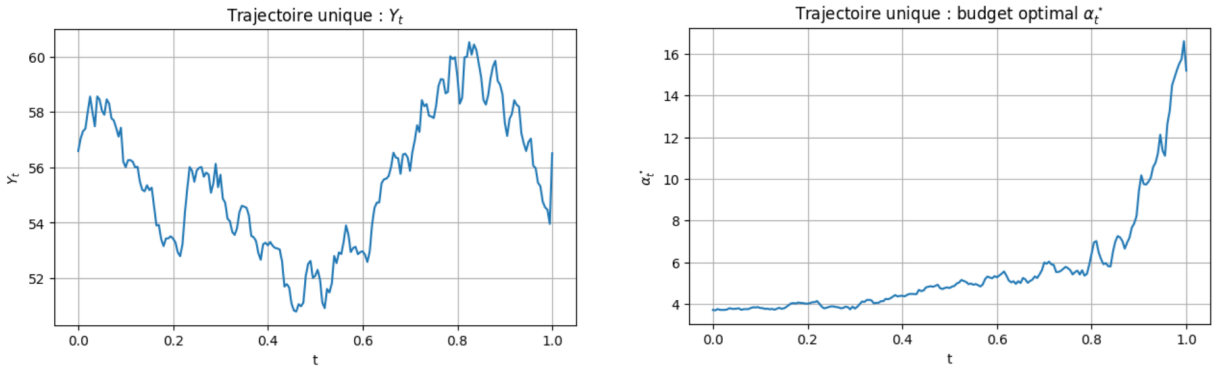


FIGURE 1.7 – Régime à forts enjeux ($S = 8$) : trajectoires de Y_t et budget optimal reconstruit $\alpha_t^\star$.

Les Figures 1.6–1.7 illustrent la transition de régime prédite par la décomposition du Hamiltonien. Le long de chaque trajectoire brownienne simulée, on reconstruit le budget instantané

optimal du Principal $\alpha_t^\star = \alpha^\star(p_t)$ en maximisant le Hamiltonien ponctuel dérivé dans la Section 1.5.3.

Dans le cas à faibles enjeux ($S = 1$), la branche sans investissement H_0 domine sur une large portion de la trajectoire, ce qui produit le plateau caractéristique $\alpha_t^\star = 1/\lambda = 2$ sous coût quadratique. En fin de trajectoire, toutefois, l'optimiseur peut basculer vers une branche d'investissement, ce qui entraîne un écart visible par rapport à ce plateau.

Dans le cas à forts enjeux ($S = 8$), une branche d'investissement domine pendant la majeure partie de la trajectoire, ce qui induit un budget fortement variable dans le temps et dépendant de l'état.

Ces observations sont cohérentes avec la structure théorique : dans le régime H_0, le maximiseur en α est constant sous coût quadratique, tandis que dans un régime d'investissement, le budget optimal est gouverné par la pente effective dans le Hamiltonien et devient très sensible à la trajectoire réalisée via $p_t = Z_t$.

1.6 Contribution : Une perspective à population infinie

1.6.1 Introduction du problème de champ moyen du Colonel Blotto

Malgré la richesse des généralisations du jeu de Blotto à des contextes asymétriques, stochastiques, discrets ou multi-joueurs, la littérature existante demeure restreinte à des modèles à nombre fini de joueurs. Or, l'étude d'une version à infinité de joueurs se justifie naturellement dès lors que l'on cherche à modéliser des situations où chaque agent affronte une masse d'adversaires nombreux, de sorte que l'interaction pertinente s'effectue avec la distribution agrégée des stratégies plutôt qu'avec chaque joueur individuellement. Par exemple, le jeu de Blotto étant un jeu de classement, il se porte particulièrement bien à ce cadre puisqu'on pourrait vouloir se classer par rapport à une foule. À ce jour, aucune contribution ne semble explorer cette approche dans le cadre des jeux à population infinie ou à champ moyen (*Mean Field Games*, MFG). Bien que la terminologie Mean Field soit plutôt utilisée pour désigner les jeux dynamiques, on utilisera cette désignation dans ce manuscrit pour le problème statique à population infinie.

L'approche à champ moyen, formalisée indépendamment par Lasry et Lions [53] et Huang, Malhamé et Caines [46], a pour but d'approximer les comportements stratégiques dans des jeux à N joueurs symétriques lorsque N tend vers l'infini. Chaque joueur interagit alors non plus avec des adversaires individualisés, mais avec une distribution agrégée représentant la population. Cela permet de transformer une situation complexe comportant N problèmes d'optimisation couplés à un seul problème d'optimisation ainsi qu'un problème de point fixe, ces derniers sur un ensemble de mesures plutôt que d'actions. Dans la suite, on donnera un cadre mathématique pour introduire le jeu à population infinie de Blotto.
Le concept d'équilibre adopté dans ce manuscrit est un équilibre statique de type Cournot–Nash / non atomique, étroitement lié aux modèles d'interactions de champ moyen. Des résultats d'existence ont été obtenus pour diverses fonctions de gain, avec information parfaite ou imparfaite ; voir par exemple Schmeidler [71] et Khan–Rustichini [2]. Il existe également un lien étroit avec les problèmes de transport optimal, comme le montrent les travaux de Blanchet et Carlier [9], [10].

Dans ce manuscrit, on utilise le terme "champ moyen" ou "Mean-Field" au sens grande population (non atomique) : les agents n'interagissent qu'à travers la distribution de la population. Le modèle est statique (one-shot), et non dynamique.

On considère un ensemble d'actions admissibles $A \subset \mathbb{R}^m$, et une fonction de gain $J(a, \mu)$ représentant ce que gagne un joueur adoptant la stratégie $a \in A$ face à une distribution de population $\mu \in \mathcal{P}(A)$, où $\mathcal{P}(A)$ est l'espace des mesures de probabilité sur A. Divers exemples de critères J sont donnés et expliqués dans le Chapitre 4.

Définition 1.6.1. *Un* équilibre de champ moyen *est alors une mesure* $\mu^* \in \mathcal{P}(A)$ *telle que :*

$$G(\mu^*) = \mu^*$$

où

$$G(\nu) : \arg \max_{\mu \in \mathcal{P}(A)} \mathbb{E}_{U \sim \mu} \left[J(U, \nu) \right]$$

Cette définition décrit que μ^* est une mesure qui représente une population jouant de manière optimale contre elle-même. De manière équivalente, μ^* peut également être définie comme suit :

$$\mathrm{Supp}(\mu^*) \subset \arg \max_{a \in A} J(a, \mu^*)$$

c'est-à-dire que tout point dans le support de μ^* est une stratégie optimale contre la distribution μ^* elle-même. On appelle ce type d'équilibres un équilibre de **Cournot-Nash**.

Une extension naturelle du jeu de Blotto dans un cadre à champ moyen consiste à considérer m champs de bataille et une population infinie de joueurs allouant des ressources sur ceux-ci. On note $S_{m-1} := \{(x_1, \dots, x_m) \in [0,1]^m \mid \sum_{i=1}^m x_i = 1\}$ le simplexe standard de dimension $m-1$, représentant les stratégies admissibles de répartition budgétaire. Soit $\mathcal{L} := \mathcal{P}(S_{m-1})$ l'ensemble des lois de probabilité sur ce simplexe.

Chaque joueur fait alors face au problème d'optimisation suivant :

$$\sup_{\nu \in \mathcal{L}} \int_{S_{m-1}} \left[\sum_{i=1}^m g_i \left(F_{\mu_i}(x_i) \right) \right] d\nu(x), \tag{1.1}$$

où $\mu \in \mathcal{L}$ désigne la distribution agrégée de la population, et ν la réponse stratégique du joueur. Le jeu de Blotto étant fondamentalement un jeu de classement, cette formulation repose sur une comparaison implicite de chaque joueur à l'ensemble de la population. Cette comparaison est rendue possible par l'introduction des fonctions de répartition marginales F_{μ_i} associées à chaque champ i, qui mesurent la proportion d'adversaires allouant moins de ressources sur ce champ.

Les fonctions $g_i : [0,1] \to \mathbb{R}$ sont supposées croissantes, et traduisent l'importance relative de se positionner favorablement sur chaque champ. Elles peuvent par exemple modéliser la valeur stratégique ou économique attachée à un bon classement sur le champ i.

Ce cadre permet d'étudier l'existence d'équilibres de champ moyen dans un contexte où les joueurs optimisent leur position relative dans une population dense, et non plus simplement leur score absolu dans un jeu à deux.

Il existe cependant des équilibres triviaux : Les masses de Dirac. En effet, toutes les masses de Dirac correspondent clairement au gain maximal, puisqu'elles permettent d'avoir une fonction de répartition presque sûrement égale à 1.

On peut alors se demander comment on pourrait trouver des équilibres non triviaux pour ce jeu. En particulier, des équilibres continus par rapport à la mesure de Lebesgue, qu'on appellera équilibres **diffus**.

1.6.2 Un manque de structure

En réalite, le problème est trop général et manque de structure. Pour s'en convaincre, on peut montrer l'existence, pour certains cas de figure, des équilibres triviaux sous forme de Diracs mais aussi d'autres diffus pour le même problème.

Par exemple, soit π une mesure absolument continue par rapport à la mesure de Lebesgue. On considère le problème avec $g_i \equiv F_{\pi_i}^{-1}$ pour tout $1 \leq i \leq m$ où $F_{\pi_i}^{-1}$ est l'inverse généralisée de la fonction de répartition de π_i : $F_{\pi_i}^{-1}(x) := \inf\{y \in \mathbb{R} | F_{\pi_i}(y) \geq x\}$.
Pour une mesure de population μ sur S_{m-1}, on rappelle que le problème de meilleure réponse consiste à maximiser

$$\nu \longmapsto \int_{S_{m-1}} \sum_{i=1}^m g_i\big(F_{\mu_i}(x_i)\big)\, d\nu(x), \tag{1.2}$$

et qu'un équilibre de champ moyen est une mesure μ telle que μ appartienne à l'ensemble des meilleures réponses à μ.

Dans le cas présent, le choix $g_i \equiv F_{\pi_i}^{-1}$ fige complètement la structure du jeu, tandis que la mesure de population μ reste libre. On vérifie d'abord que π elle-même est un équilibre diffus pour ce jeu. En effet, en prenant $\mu = \pi$, on obtient, pour tout $x \in S_{m-1}$,

$$g_i\big(F_{\pi_i}(x_i)\big) = F_{\pi_i}^{-1}\big(F_{\pi_i}(x_i)\big) = x_i,$$

où l'on a utilisé que π_i est absolument continue (et, par exemple, à densité strictement positive) de sorte que F_{π_i} est continue et strictement croissante, donc inversible. Par conséquent,

$$\sum_{i=1}^{m} g_i\big(F_{\pi_i}(x_i)\big) = \sum_{i=1}^{m} x_i = 1 \quad \text{pour tout } x \in S_{m-1},$$

et la fonctionnelle de meilleure réponse s'écrit

$$\int_{S_{m-1}} \sum_{i=1}^{m} g_i\big(F_{\pi_i}(x_i)\big)\, d\nu(x) = \int_{S_{m-1}} 1\, d\nu(x) = 1,$$

indépendamment du choix de ν. Autrement dit, *toute* mesure ν est une meilleure réponse à π, et en particulier π est un équilibre de champ moyen diffus, de valeur d'équilibre égale à 1.

D'un autre côté, le même choix des fonctions $(g_i)_{1\le i\le m}$ engendre aussi des équilibres purement atomiques. Fixons un point $x^\star = (x_1^\star, \dots, x_m^\star) \in S_{m-1}$ tel que $x_i^\star > 0$ pour tout i, et considérons la mesure de population $\mu = \delta_{x^\star}$. Pour tout i, la marginale μ_i est alors la Dirac en $x_i^\star$, et sa fonction de répartition vaut

$$F_{\mu_i}(y) = \begin{cases} 0, & y < x_i^\star, \\ 1, & y \ge x_i^\star. \end{cases}$$

On en déduit que, pour tout $x \in S_{m-1}$,

$$g_i\big(F_{\mu_i}(x_i)\big) = \begin{cases} F_{\pi_i}^{-1}(0), & x_i < x_i^\star, \\ F_{\pi_i}^{-1}(1), & x_i \ge x_i^\star. \end{cases}$$

En particulier, au point $x = x^\star$ on a $F_{\mu_i}(x_i^\star) = 1$ pour tout i, donc

$$\sum_{i=1}^{m} g_i\big(F_{\mu_i}(x_i^\star)\big) = \sum_{i=1}^{m} F_{\pi_i}^{-1}(1).$$

Comme chaque terme $g_i(F_{\mu_i}(x_i))$ est borné par $F_{\pi_i}^{-1}(1)$, la somme ci-dessus est un maximum, atteint uniquement lorsque $x = x^\star$ (du fait de la contrainte $\sum_i x_i = 1$ et de la positivité des $x_i^\star$). Par conséquent, la meilleure réponse à la population $\mu = \delta_{x^\star}$ est précisément de jouer $x^\star$, et $\delta_{x^\star}$ est donc aussi un équilibre de champ moyen. La valeur d'équilibre associée à $\delta_{x^\star}$ vaut

$$\int_{S_{m-1}} \sum_{i=1}^{m} g_i\big(F_{\mu_i}(x_i)\big)\, d\mu(x) = \sum_{i=1}^{m} F_{\pi_i}^{-1}(1),$$

qui est typiquement strictement plus grande que 1. Par exemple, si chaque marginale π_i est supportée dans $[0,1]$ avec $F_{\pi_i}^{-1}(1) = 1$, on obtient une valeur d'équilibre égale à m pour $\delta_{x^\star}$, tandis

que l'équilibre diffus π garde une valeur égale à 1.

Cet exemple montre qu'avec ce choix de fonctions $(g_i)_{1\leq i\leq m}$ le même jeu admet simultanément des équilibres de champ moyen diffus et des équilibres purement atomiques, avec des valeurs d'équilibre différentes. En particulier, il n'existe pas de garanties d'unicité ou même de régularité de l'équilibre. Cela illustre que, dans toute sa généralité, le problème considéré est *trop* flexible et ne possède pas la structure supplémentaire nécessaire pour obtenir un équilibre diffus bien déterminé.

On se propose alors dans la section suivante d'essayer de rajouter des conditions sous forme de pénalisation pour pouvoir traiter une version du problème un peu plus intéressante.

1.6.3 Un modèle-jouet de champ moyen

Avant de passer à la version du problème qu'on a élu d'étudier, on se propose dans un premier temps d'étudier une version simplifiée, ou un "modèle-jouet" du problème. Dans cette version, on se restreint à un ensemble de mesures absolument continues par rapport à la mesure de Lebesgue réduit, de sorte qu'on puisse avoir une compacité ainsi qu'une grande régularité des fonctions de densité. On prouve dans ce contexte restreint, à l'aide d'outils assez intéressants, l'existence d'un équilibre à champ moyen. Ce modèle-jouet ainsi que ses résultats sont explicités en Annexe B

1.6.4 Le jeu d'allocation à champ moyen avec pénalisation entropique

On choisit de travailler avec une régularisation entropique. Ce choix est inspiré par des développements en transport optimal et en optimisation de réseaux de neurones. En transport optimal, on optimise une fonctionnelle linéaire sur un espace de mesures de probabilité ; l'ajout d'un terme entropique rend alors l'objectif strictement convexe (voir [7]). De même, en optimisation de réseaux de neurones, un problème non convexe sur les paramètres du réseau peut être relaxé en un problème linéaire sur des mesures de probabilité, puis rendu strictement convexe via une régularisation entropique (voir [18]). Une telle pénalisation est donc connue pour bien se comporter théoriquement et numériquement.

Dans cette section, on fixe comme mesure de référence λ la mesure de Lebesgue sur le simplexe S_{m-1}. On considère alors, pour $\beta > 0$, le problème régularisé

$$\sup_{\nu\in\mathcal{L}} \left\{ \int_{S_{m-1}} R(x,\mu)\,\nu(\mathrm{d}x) \;-\; \frac{1}{\beta}\,H(\nu \mid \lambda) \right\}, \tag{1.3}$$

où $H(\nu \mid \lambda)$ désigne l'entropie relative, ou encore la divergence de Kullback–Leibler,

$$H(\nu \mid \lambda) = \begin{cases} \displaystyle\int_{S_{m-1}} f_\nu(x)\log(f_\nu(x))d\lambda(x), & \text{si } \nu \ll \lambda, \\ +\infty, & \text{sinon.} \end{cases}$$

L'expression (1.3) est la forme classique d'un principe d'énergie libre. Le terme $\int R\,\mathrm{d}\nu$ joue le rôle d'une énergie (ou d'une utilité moyenne), tandis que $H(\nu \mid \lambda)$ mesure une entropie relative, i.e. un coût à s'écarter d'un "bruit" ou d'une référence. Le paramètre β s'interprète comme une température inverse : à température élevée (petit β), la distribution optimale reste diffuse et proche de la référence ; à température basse (grand β), elle se concentre davantage autour

des allocations qui donnent un gain élevé. Dans des problèmes d'allocation de ressources, cette lecture est naturelle : les choix agrègent des facteurs non modélisés (fluctuations, imprécisions, hétérogénéité des joueurs, etc.) dont l'intensité se résume par la température.

Le terme $H(\nu \mid \lambda)$ peut aussi se lire comme un coût de déformation de la mesure de Lebesgue : passer d'une loi "neutre" à une loi qui présente des pics, ou qui concentre une part importante de sa masse sur des régions fines du simplexe, devient coûteux. Autrement dit, la pénalisation exprime que rendre une stratégie très pointue n'est pas gratuit : elle exige de réduire fortement la dispersion, ce qui, dans des contextes applicatifs, correspond à une exigence d'information, d'organisation, de coordination ou de précision d'exécution. Le fait que $H(\nu \mid \lambda) = +\infty$ dès que ν n'est pas absolument continue par rapport à Lebesgue formalise l'idée qu'une stratégie purement atomique (ou plus généralement singulière) est vue comme une limite idéalisée, infiniment coûteuse relativement à une référence continue.

Du point de vue de la théorie de l'information, $H(\nu \mid \lambda)$ quantifie aussi la complexité de décision. Déformer Lebesgue pour privilégier certaines allocations revient à injecter de l'information (ou à effectuer un effort de calcul) afin de distinguer finement des choix dans un espace continu. Dans cette lecture, l'agent ne suit pas une rationalité parfaite : il arbitre entre performance et coût d'information. Le paramètre β traduit alors une capacité à s'écarter de la stratégie de référence. Cette interprétation est équivalente à une formulation contrainte : maximiser le gain moyen sous une contrainte $H(\nu \mid \lambda) \leq \kappa$, et passer à la forme (1.3) via un multiplicateur de Lagrange. Le paramètre β joue le rôle de ce multiplicateur et paramétrise un ensemble de régimes, de très diffus à très concentré.

Il est important de souligner que cette régularisation ne dépend pas de la forme particulière du gain. Dans la section précédente de ce manuscrit, on étudiait un payoff qui s'écrit :

$$\sup_{\nu \in \mathcal{P}(S)} \int_S \Big[\sum_{i=1}^{m} g_i\big(F_{\mu_i}(x_i)\big) \Big] \, \nu(\mathrm{d}x),$$

où les indices correspondent aux marginales. Cette structure correspond à un choix spécifique de $R(x, \mu) = \sum_{i=1}^{m} g_i(F_{\mu_i}(x_i))$. La pénalisation entropique permet de s'affranchir de cette restriction sans changer le formalisme : on peut, dans cette sous-section, se permettre de discuter un gain général $R : S_{m-1} \times \mathcal{L} \to \mathbb{R}$, qu'il soit additif, non additif, spectral, ou qu'il combine plusieurs effets. L'entropie joue alors le rôle d'une brique de modélisation transversale, commune à une famille plus large de jeux d'allocation.

Enfin, cette pénalisation a un avantage formel : ajouter $-\frac{1}{\beta} H(\nu \mid \lambda)$ rend le critère strictement concave en ν, ce qui sélectionne une meilleure réponse unique. De plus, la solution prend une forme explicite de type mesure de Gibbs, qui sera explicitée plus tard dans cette section.

Meilleure réponse entropique

On considère, pour un gain général $R : S_{m-1} \times \mathcal{L} \to \mathbb{R}$, le problème de meilleure réponse à une population μ :

$$\sup_{\nu \in \mathcal{L}} \left\{ \int_{S_{m-1}} R(x, \mu) \, \nu(\mathrm{d}x) \; - \; \frac{1}{\beta} \, \mathrm{KL}(\nu, \lambda) \right\}, \qquad \beta > 0. \tag{1.4}$$

Lemme 1.6.2. *On suppose que $x \mapsto R(x, \mu)$ soit bornée et mesurable pour tout μ. Alors le problème* (1.4) *admet un unique maximiseur $\nu = \mathrm{BR}(\mu)$, donné par*

$$\mathrm{BR}(\mu)(\mathrm{d}x) = \frac{\exp(\beta R(x, \mu))}{Z(\mu)} \, \lambda(\mathrm{d}x), \qquad Z(\mu) = \int_{S_{m-1}} \exp(\beta R(y, \mu)) \, \lambda(\mathrm{d}y). \tag{1.5}$$

En particulier, $\mathrm{BR}(\mu) \ll \lambda$ *et sa densité vaut* $f_{\mathrm{BR}(\mu)}(x) = \exp(\beta R(x,\mu))/Z(\mu)$.

La preuve est explicitée dans le chapitre 4.

Remarque 1.6.3. *La meilleure réponse est un objet explicite et absolument continu par rapport à* λ*, ce qui impose en particulier que la stratégie optimale ne crée pas de masses atomiques.*

Équilibre de champ moyen général

On montre dans cette partie l'existence d'un équilibre de champ moyen avec peu d'hypothèses sur R.

Définition 1.6.4. *On définit l'opérateur de meilleure réponse*

$$T : \mathcal{L} \to \mathcal{L}, \qquad T(\mu) = \mathrm{BR}(\mu),$$

où $\mathrm{BR}(\mu)$ *est donné par* (1.5).

Définition 1.6.5 (Équilibre de champ moyen)**.** *Un* équilibre de champ moyen *(MFE) est un point fixe* $\mu^\star \in \mathcal{L}$ *tel que*

$$\mu^\star = T(\mu^\star).$$

Hypothèse 1.6.6. *On suppose :*

*(A1) (*Bornitude*) il existe* $C < \infty$ *tel que* $|R(x,\mu)| \le C$ *pour tout* $x \in S_{m-1}$ *et tout* $\mu \in \mathcal{L}$ *;*

*(A2) (*Continuité en μ*) si* $\mu_n \Rightarrow \mu$ *faiblement, alors* $R(x,\mu_n) \to R(x,\mu)$*,* λ*-presque partout.*

Théorème 1.6.7 (Existence d'un MFE)**.** *Sous l'Hypothèse 1.6.6, l'application* T *est continue pour la convergence faible sur* $\mathcal{P}(S_{m-1})$*. En particulier, elle admet au moins un point fixe* $\mu^\star$*, donc un MFE existe.*

La preuve est explicitée dans le chapitre 4.

Remarque 1.6.8. *Dans la partie précédente de la thèse, le gain considéré était de la forme*

$$R(x,\mu) = \sum_{i=1}^{m} g_i\big(F_{\mu_i}(x_i)\big),$$

où μ_i *désigne la* i*-ième marginale de* μ*. Ce cas s'inscrit bien dans le formalisme ci-dessus : l'opérateur* T *reste donné par* (1.5) *avec ce choix de* R*. L'intérêt de présenter d'abord le cas général est qu'on peut ensuite changer* R *(non additif, spectral, avec effets de complémentarité, etc.) sans toucher à la structure de la meilleure réponse, ni à la définition de l'équilibre.*

1.6.5 Construction d'un équilibre à champ moyen du jeu de Blotto

On s'intéresse à la formulation du jeu d'allocation étudié qui s'apparente le plus au jeu de Blotto :

$$R_\alpha(x,\mu) = \sum_{i=1}^{m} \mathbf{1}_{\{F_{\mu_i}(x_i) \ge \alpha\}} = \sum_{i=1}^{m} \mathbf{1}_{\{x_i \ge q_{\mu_i}(\alpha)\}}, \qquad \alpha \in (0,1).$$

Ce choix de formulation en quantiles est particulièrement naturel dans Blotto, car le jeu est fondamentalement un jeu de classement. Sur chaque champ de bataille i, l'agent ne tire pas un gain de la valeur absolue x_i, mais de sa position relative dans la population : il "gagne" le champ

i s'il fait partie des α meilleurs (ou, de manière équivalente, s'il est dans le top $(1-\alpha)$) sur cette coordonnée. Lorsque la population joue selon μ, cette condition se lit exactement via la fonction de répartition marginale F_{μ_i} : avoir $F_{\mu_i}(x_i) \geq \alpha$ signifie que la proportion d'agents ayant une allocation *inférieure ou égale* à x_i sur le champ i est au moins α, donc que x_i dépasse le seuil correspondant au α-quantile $q_{\mu_i}(\alpha)$. On obtient ainsi l'écriture

$$\mathbf{1}_{\{F_{\mu_i}(x_i)\geq\alpha\}} = \mathbf{1}_{\{x_i\geq q_{\mu_i}(\alpha)\}},$$

qui rend explicite le fait que, à population fixée, le gain de Blotto est un payoff à seuil déterminé endogènement par les quantiles de la distribution de la population.

Dans le cas homogène (champs de bataille équivalents), on cherche un équilibre de champ moyen dans la classe des lois invariantes par permutation des coordonnées. Cette symétrie reflète le fait que ni la contrainte $x \in S_{m-1}$ ni le payoff de Blotto ne distinguent a priori un champ de bataille plutôt qu'un autre.

Définition 1.6.9. *Une mesure $\mu \in \mathcal{L}$ est dite* échangeable *si, pour toute permutation σ de $\{1,\dots,m\}$, l'image de μ par l'application $x \mapsto (x_{\sigma(1)},\dots,x_{\sigma(m)})$ coïncide avec μ. Dans ce cas, toutes les marginales $\mu_1,\dots,\mu_m$ sont identiques ; on note $\mu_{(1)}$ la marginale commune et F_μ sa fonction de répartition.*

Pour le payoff de Blotto et une population μ échangeable, les seuils de victoire sur chaque champ de bataille coïncident : $q_{\mu_i}(\alpha) = q_\mu(\alpha)$ pour tout i, où $q_\mu(\alpha)$ est le α-quantile de la marginale commune.

Définition 1.6.10. *Pour $q \in [0,1]$, on définit une mesure $\mu_q \in \mathcal{L}$ par*

$$\mu_q(\mathrm{d}x) = \frac{\exp\big(\beta\, s_q(x)\big)}{Z_m(q)}\,\lambda(\mathrm{d}x), \qquad s_q(x) = \sum_{i=1}^{m} \mathbf{1}_{\{x_i\geq q\}}, \qquad Z_m(q) = \int_{S_{m-1}} \exp\big(\beta\, s_q(y)\big)\,\lambda(\mathrm{d}y). \tag{1.6}$$

La mesure μ_q est échangeable ; on note $F_{(q)}$ la fonction de répartition de sa marginale commune.

L'idée est que, dans un équilibre échangeable, le seuil de victoire doit être le même sur chaque champ de bataille, donc résumable par un scalaire q. On impose ensuite la cohérence : le seuil q doit coïncider avec le α-quantile de la marginale qu'il engendre.

Soit à présent $\alpha \in (0,1)$ et une population $\mu_{q^\star}$. Par définition de Blotto,

$$R_\alpha(x,\mu_{q^\star}) = \sum_{i=1}^{m} \mathbf{1}_{\{F_{\mu_{q^\star,i}}(x_i)\geq\alpha\}},$$

où $\mu_{q^\star,i}$ est la i-ième marginale de $\mu_{q^\star}$. Comme $\mu_{q^\star}$ est échangeable, toutes ses marginales coïncident et ont pour CDF $F_{(q^\star)}$. La condition (1.7) signifie que le α-quantile de cette marginale est $q^\star$, donc, pour tout i,

$$\mathbf{1}_{\{F_{\mu_{q^\star,i}}(x_i)\geq\alpha\}} = \mathbf{1}_{\{x_i\geq q^\star\}}.$$

Ainsi,

$$R_\alpha(x,\mu_{q^\star}) = \sum_{i=1}^{m} \mathbf{1}_{\{x_i\geq q^\star\}} = s_{q^\star}(x).$$

La meilleure réponse entropique à $\mu_{q^\star}$, d'après le lemme 1.6.2, est donc la mesure de densité proportionnelle à $\exp(\beta R_\alpha(x, \mu_{q^\star})) = \exp(\beta s_{q^\star}(x))$ par rapport à λ, c'est-à-dire exactement $\mu_{q^\star}$ au sens de (1.6). On a donc

$$\mathrm{BR}(\mu_{q^\star}) = \mu_{q^\star},$$

ce qui prouve la proposition suivante :

Proposition 1.6.11 (Construction d'équilibre)**.** *Soit $\alpha \in (0,1)$. On suppose qu'il existe $q^\star \in [0,1]$ tel que*

$$F_{(q^\star)}(q^\star) = \alpha, \qquad i.e. \qquad q^\star = F^{-1}_{(q^\star)}(\alpha). \tag{1.7}$$

Alors $\mu_{q^\star}$ est un équilibre de champ moyen (MFE) pour le payoff de Blotto dans la classe des lois échangeables.

Remarque 1.6.12. *La proposition réduit la recherche d'un MFE échangeable à l'existence d'un point fixe scalaire $q^\star$ de la relation de quantile* (1.7)*. On remplace ainsi un point fixe dans l'espace infini-dimensionnel $\mathcal{L}$ par la résolution d'une équation en une variable. Dans les sections suivantes, on calcule $F_{(q)}(q)$ explicitement pour $m = 2$ et $m = 3$, ce qui donne des formules fermées (ou des équations quadratiques) pour $q^\star$.*

La preuve de la Proposition 1.6.11 repose uniquement sur le fait que, pour Blotto, le payoff est déterminé par des *seuils marginaux* (des quantiles). Cette idée se généralise sans changement conceptuel au cas non homogène, où les champs de bataille ne sont plus équivalents. Dans ce cadre, on remplace le seuil scalaire q par un vecteur $q = (q_1, \dots, q_m)$.

Définition 1.6.13. *Pour $q = (q_1, \dots, q_m) \in [0,1]^m$, on définit une mesure $\mu_q \in \mathcal{L}$ par*

$$\mu_q(\mathrm{d}x) = \frac{\exp\big(\beta\, s_q(x)\big)}{Z(q)}\,\lambda(\mathrm{d}x), \qquad s_q(x) := \sum_{i=1}^{m} \mathbf{1}_{\{x_i \geq q_i\}}, \qquad Z(q) := \int_{S_{m-1}} \exp\big(\beta\, s_q(y)\big)\,\lambda(\mathrm{d}y). \tag{1.8}$$

On note $F_i^{(q)}$ la fonction de répartition de la i-ième marginale de μ_q.

Dans la version non homogène, on peut considérer par exemple le payoff

$$R(x, \mu) = \sum_{i=1}^{m} v_i\, \mathbf{1}_{\{F_{\mu_i}(x_i) \geq \alpha_i\}},$$

où (v_i) sont des valeurs de champs de bataille et $(\alpha_i) \in (0,1)^m$ des niveaux de classement. Le point clé est que la *forme* des ensembles de gain ne dépend que des quantiles marginaux, donc d'un vecteur de seuils.

Définition 1.6.14. *On fixe $\alpha = (\alpha_1, \dots, \alpha_m) \in (0,1)^m$ et $v = (v_1, \dots, v_m) \in (0,\infty)^m$, et l'on considère le payoff*

$$R_{\alpha,v}(x, \mu) := \sum_{i=1}^{m} v_i\, \mathbf{1}_{\{F_{\mu_i}(x_i) \geq \alpha_i\}}.$$

Pour tout vecteur de seuils $q = (q_1, \dots, q_m) \in [0,1]^m$, on définit la mesure $\mu_q^v \in \mathcal{L}$ par

$$\mu_q^v(\mathrm{d}x) := \frac{1}{Z_v(q)} \exp\Big(\beta\, s_q^v(x)\Big)\,\lambda(\mathrm{d}x), \qquad s_q^v(x) := \sum_{i=1}^{m} v_i\, \mathbf{1}_{\{x_i \geq q_i\}},$$

$$Z_v(q) := \int_{S_{m-1}} \exp\Big(\beta\, s_q^v(y)\Big)\,\lambda(\mathrm{d}y).$$

On note $F_i^{(v,q)}$ la fonction de répartition de la i-ième marginale de μ_q^v.

Proposition 1.6.15 (Construction d'équilibre de Blotto général)**.** *On suppose qu'il existe* $q^\star = (q_1^\star, \dots, q_m^\star) \in [0,1]^m$ *tel que, pour tout* $i = 1, \dots, m$*,*

$$F_i^{(v,q^\star)}(q_i^\star) = \alpha_i \qquad (i.e.\ q_i^\star = (F_i^{(v,q^\star)})^{-1}(\alpha_i)). \tag{1.9}$$

Alors $\mu^v_{q^\star}$ *est un équilibre de champ moyen (MFE) pour le payoff* $R_{\alpha,v}$*.*

En particulier, la recherche d'un MFE de cette forme se ramène à la recherche d'un point fixe de l'application

$$\Psi : [0,1]^m \to [0,1]^m, \qquad \Psi(q) := \big((F_1^{(v,q)})^{-1}(\alpha_1), \dots, (F_m^{(v,q)})^{-1}(\alpha_m)\big).$$

La preuve se fait assez similaire à la proposition 1.6.11 : si on a $q^\star$ satisfaisant (1.9) et on considère la population $\mu^v_{q^\star}$. Par définition,

$$R_{\alpha,v}(x, \mu^v_{q^\star}) = \sum_{i=1}^m v_i \, \mathbf{1}_{\{F_{\mu^v_{q^\star,i}}(x_i) \geq \alpha_i\}} = \sum_{i=1}^m v_i \, \mathbf{1}_{\{F_i^{(v,q^\star)}(x_i) \geq \alpha_i\}},$$

où $\mu^v_{q^\star,i}$ désigne la i-ième marginale de $\mu^v_{q^\star}$. La condition (1.9) signifie que $q_i^\star$ est le α_i-quantile de cette marginale, donc

$$\mathbf{1}_{\{F_i^{(v,q^\star)}(x_i) \geq \alpha_i\}} = \mathbf{1}_{\{x_i \geq q_i^\star\}}, \qquad i = 1, \dots, m.$$

Ainsi,

$$R_{\alpha,v}(x, \mu^v_{q^\star}) = \sum_{i=1}^m v_i \, \mathbf{1}_{\{x_i \geq q_i^\star\}} = s^v_{q^\star}(x).$$

Par la forme de la meilleure réponse (cf. (1.5)), la meilleure réponse entropique à $\mu^v_{q^\star}$ a une densité proportionnelle à $\exp(\beta R_{\alpha,v}(x, \mu^v_{q^\star})) = \exp(\beta s^v_{q^\star}(x))$ par rapport à λ. C'est donc exactement la mesure $\mu^v_{q^\star}$ par définition, ce qui montre $\mathrm{BR}(\mu^v_{q^\star}) = \mu^v_{q^\star}$.

Remarque 1.6.16. *Dans le cas homogène (mêmes* $\alpha_i = \alpha$ *et mêmes poids* $v_i \equiv 1$*), l'échangeabilité force* $q_1 = \cdots = q_m$ *et l'on retrouve le point fixe scalaire. Dans le cas général, l'objet naturel est un point fixe vectoriel, et l'intérêt principal est conceptuel : on remplace un point fixe dans l'espace des mesures par un point fixe sur* $[0,1]^m$*.*

Remarque 1.6.17. *L'existence d'un vecteur de seuils* $q^\star \in [0,1]^m$ *satisfaisant les conditions de quantiles (une par marginale) est automatique. En effet, l'existence d'un MFE* $\mu^\star$ *est déjà garantie par le théorème d'existence général. À partir de cet équilibre, il suffit de définir* $q_i^\star$ *comme le* α_i*-quantile de la* i*-ième marginale de* $\mu^\star$*. Par construction, ces seuils reproduisent exactement les conditions de classement qui entrent dans le payoff de Blotto, et ils sont donc cohérents avec la meilleure réponse.*

Équilibres explicites du jeu de Blotto pénalisé en 2 dimensions

Avec les résultats de cette section, on peut à présent donner explicitement la forme analytiques des équilibres à champ moyen en petites dimensions, on commence par le cas $m = 2$:

Proposition 1.6.18. *On fixe* $\alpha \in (0,1)$ *et on pose* $a := e^\beta > 1$*. À l'équilibre de champ moyen (échangeable), on tire*

$$(X, 1-X) \in S_1, \qquad X \in [0,1],$$

où X admet une densité (par rapport à la Lebesgue sur $[0,1]$) donnée par

$$f^\star(x) = \begin{cases} \dfrac{1+2\alpha(a-1)}{a}, & 0 \le x < \dfrac{a\alpha}{1+2\alpha(a-1)}, \\ 1+2\alpha(a-1), & \dfrac{a\alpha}{1+2\alpha(a-1)} \le x \le 1 - \dfrac{a\alpha}{1+2\alpha(a-1)}, \\ \dfrac{1+2\alpha(a-1)}{a}, & 1 - \dfrac{a\alpha}{1+2\alpha(a-1)} < x \le 1, \end{cases} \qquad \textit{si } \alpha \le \frac{1}{2},$$

et

$$f^\star(x) = \begin{cases} a(2\alpha-1)+2(1-\alpha), & 0 \le x \le \dfrac{1-\alpha}{a(2\alpha-1)+2(1-\alpha)}, \\ \dfrac{a(2\alpha-1)+2(1-\alpha)}{a}, & \dfrac{1-\alpha}{a(2\alpha-1)+2(1-\alpha)} < x < \dfrac{a(2\alpha-1)+(1-\alpha)}{a(2\alpha-1)+2(1-\alpha)}, \\ a(2\alpha-1)+2(1-\alpha), & \dfrac{a(2\alpha-1)+(1-\alpha)}{a(2\alpha-1)+2(1-\alpha)} \le x \le 1, \end{cases} \qquad \textit{sinon.}$$

Une description plus détaillée du résultat est présente au Chapitre 4, mais on peut déjà remarquer en dimension 2 que la Proposition précédente fait apparaître deux régimes séparés par le seuil critique $\alpha = \frac{1}{2}$. L'équilibre est décrit par une densité constante par morceaux, déterminée par un seuil endogène $q^\star(\alpha,\beta)$ (quantile marginal sous μ_q), qui découpe $[0,1]$ en zones correspondant à deux, une, ou zéro victoire.

Le paramètre β ne change pas quelles régions sont favorisées, mais seulement l'intensité de la concentration : le rapport entre plateaux vaut $\exp(\beta)$. Quand $\beta \to 0$, l'équilibre tend vers la loi uniforme (et $q^\star \to \alpha$). Quand $\beta \to \infty$, l'équilibre se concentre sur les allocations maximisant le nombre de victoires : près de $X = \frac{1}{2}$ si $\alpha < \frac{1}{2}$, et près des coins si $\alpha > \frac{1}{2}$. Enfin, faire croître α augmente le seuil $q^\star$ et rend les victoires plus difficiles, ce qui contracte la région « deux victoires » lorsque $\alpha \uparrow \frac{1}{2}$ et élargit la région « zéro victoire » lorsque $\alpha > \frac{1}{2}$.

Équilibres explicites du jeu de Blotto pénalisé en 3 dimensions

Pour le cas $m = 3$, l'équilibre se calcule encore explicitement, on commence par définir les seuils suivants :

Définition 1.6.19. *On fixe $\alpha \in (0,1)$ et on pose $a := e^\beta > 1$. On définit*

$$\alpha_{1/3} := \frac{a+4}{3(a+2)}, \qquad \alpha_{1/2} := \frac{2a+1}{3a+1}.$$

On définit ensuite $t = t(\alpha) \in (0,1)$ comme l'unique solution de l'équation quadratique suivante :

$$\begin{cases} A_1 t^2 + B_1 t + C_1 = 0, & \textit{si } \alpha \le \alpha_{1/3} \quad (\textit{avec } t \in (0, \frac{1}{3}]), \\ A_2 t^2 + B_2 t + C_2 = 0, & \textit{si } \alpha_{1/3} < \alpha \le \alpha_{1/2} \quad (\textit{avec } t \in (\frac{1}{3}, \frac{1}{2}]), \\ A_3 t^2 + B_3 t + C_3 = 0, & \textit{si } \alpha > \alpha_{1/2} \quad (\textit{avec } t \in (\frac{1}{2}, 1)), \end{cases}$$

où

$$A_1 = 4 - 5a + \alpha(15a - 9a^2 - 6), \quad B_1 = 2a + \alpha(6a^2 - 6a), \quad C_1 = -\alpha a^2,$$
$$A_2 = 4a^2 - 14a + 9 + \alpha(-12a^2 + 21a - 9),$$
$$B_2 = -4a^2 + 12a - 6 + \alpha(12a^2 - 18a + 6), \quad C_2 = (a-1)^2 + \alpha(-3a^2 + 3a - 1),$$
$$A_3 = 2a - 3 + \alpha(-3a + 3), \quad B_3 = -4a + 6 + \alpha(6a - 6), \quad C_3 = 2a - 2 + \alpha(-3a + 2).$$

Définition 1.6.20. *Pour $t \in (0,1)$, on pose, pour $x = (x_1, x_2, x_3) \in S_2$,*

$$k_t(x) := \#\{i \in \{1,2,3\} : \ x_i \geq t\}.$$

On définit la densité (par rapport à la Lebesgue sur S_2)

$$f_t(x) := \frac{a^{k_t(x)}}{Z(t)},$$

où $Z(t)$ est donné explicitement par

$$Z(t) = \begin{cases} aA_1(t) + a^2A_2(t) + a^3A_3(t), & t \leq \frac{1}{3}, \\ A_0(t) + aA_1(t) + a^2A_2(t), & \frac{1}{3} < t \leq \frac{1}{2}, \\ A_0(t) + aA_1(t), & t > \frac{1}{2}, \end{cases}$$

avec

$$\text{si } t \leq \frac{1}{3} : \quad A_3(t) = \frac{(1-3t)^2}{2}, \qquad A_2(t) = \frac{3\big((1-2t)^2 - (1-3t)^2\big)}{2}, \qquad A_1(t) = \frac{1}{2} - A_2(t) - A_3(t),$$

$$\text{si } \frac{1}{3} < t \leq \frac{1}{2} : \quad A_2(t) = \frac{3(1-2t)^2}{2}, \qquad A_0(t) = \frac{(3t-1)^2}{2}, \qquad A_1(t) = \frac{1}{2} - A_0(t) - A_2(t),$$

$$\text{si } t > \frac{1}{2} : \quad A_1(t) = \frac{3(1-t)^2}{2}, \qquad A_0(t) = \frac{1}{2} - A_1(t).$$

On note μ_t la loi sur S_2 de densité f_t.

Proposition 1.6.21. *On fixe $\alpha \in (0,1)$, on définit $t = t(\alpha)$ comme dans la Définition 1.6.19, et on pose $\mu^\star := \mu_{t(\alpha)}$ comme dans la Définition 1.6.20. Alors $\mu^\star$ est un MFE (échangeable) pour le payoff homogène*

$$R_\alpha(x, \mu) = \sum_{i=1}^{3} \mathbf{1}_{\{F_{\mu_i}(x_i) \geq \alpha\}}.$$

En particulier, la densité d'équilibre est donnée par

$$f^\star(x) = \frac{a^{k_{t(\alpha)}(x)}}{Z(t(\alpha))}, \qquad x \in S_2,$$

La preuve est détaillée dans le Chapitre 4.

En dimension 3, l'équilibre conserve la même structure de base qu'en dimension 2 : la densité d'équilibre est constante par morceaux et favorise les allocations qui remportent davantage de champs, avec un facteur multiplicatif $a = e^\beta$ par victoire supplémentaire. La géométrie du simplexe introduit toutefois trois régimes distincts, séparés par les niveaux critiques $\alpha_{1/3}$ et $\alpha_{1/2}$. L'équilibre est paramétré par le seuil $t = t(\alpha)$ et la loi de Gibbs μ_t ; augmenter β transfère simplement de la masse vers les régions où l'on gagne plus de champs.

La transition entre régimes s'explique par des contraintes de faisabilité : si $t \leq \frac{1}{3}$, on peut avoir trois coordonnées au-dessus de t (victoire triple) ; si $\frac{1}{3} < t \leq \frac{1}{2}$, les victoires triples deviennent impossibles mais les victoires doubles restent possibles ; si $t > \frac{1}{2}$, même les victoires doubles sont impossibles et l'on ne peut gagner qu'un seul champ. Ainsi, via $t(\alpha)$, le paramètre α sélectionne les motifs de victoire géométriquement accessibles.

Comme en dimension 2, lorsque $\beta \to 0$ (donc $a \to 1$) l'équilibre devient proche de l'uniforme sur S_2 ; à l'inverse, lorsque $\beta \to \infty$ (donc $a \to \infty$), la mesure se concentre sur les régions où le nombre de victoires est maximal dans le régime considéré : près du barycentre lorsque la victoire triple est possible, près de profils "deux grands–un petit" lorsque seules les doubles victoires sont possibles, et près des sommets lorsque seule une victoire est possible.

Enfin, il est utile de distinguer ce qui est imposé par la contrainte de quantile de ce qui dépend de β. Par construction de $t(\alpha)$, on a $F_{(t(\alpha))}(t(\alpha)) = \alpha$, d'où

$$\mathbb{P}(X_i \geq t(\alpha)) = 1 - \alpha,$$

et donc un nombre moyen de victoires égal à $3(1-\alpha)$, indépendamment de β. En revanche, β modifie la répartition du nombre de victoires (gagner $0, 1, 2$ ou 3 champs), via les poids relatifs $a^{k_t(x)}$.

Une description plus détaillée est présente dans le Chapitre 4

Chapitre 2

Stochastic differential Blotto game

2.1 Introduction

Colonel Blotto games are a class of resource allocation games, i.e. games where each player must distribute a finite budget across multiple battlefields while competing against the other players. In Blotto games, the goal of each player is to maximize an "all-or-nothing" type criterion, in the sense that, on a given battlefield, the player who has allocated the most resources wins the entire value of that battlefield. More generally, we can consider the case of a payoff function that depends on the difference in the players' state processes on each battlefield. In this chapter, we consider the case of two-player resource allocation games with a very general payoff function. These models have a wide range of potential applications : in cybersecurity [34], telecommunication models [35, 33], and economic competition [51].

Despite their description and analysis by Borel [12] already in 1921, equilibrium distributions for Colonel Blotto (CB) games have been characterized in a satisfying way only recently.

In this chapter, we introduce a continuous-time version of ressource allocation games. The mathematical framework of the game is rigorously introduced in Section 3.2., but we already informally describe the game here. Let $m \geq 1$ be the number of battlefields of the game. If we denote X^1 and X^2 the m-dimensional state processes of players 1 and 2, we are interested in games where player 1 aims to maximize the quantity $\mathbb{E}\left[g(X_T^1 - X_T^2)\right]$, while player to aims to minimize it, where $g : \mathbb{R}^m \to \mathbb{R}$ is a gain function. We consider the controlled state dynamics

$$dX_t^1 = a_t\, dt + dW_t^1, \qquad dX_t^2 = b_t\, dt + dW_t^2,$$

where W^1 and W^2 are independent m-dimensional Brownian motions. Player 1 controls the drift $a = (a_t)_{0\leq t\leq T}$ and Player 2 controls the drift $b = (b_t)_{0\leq t\leq T}$. The budget constraints are encoded by the admissible sets

$$A := \Big\{x \in \mathbb{R}^m :\ x_i \geq 0 \text{ for all } i,\ \sum_{i=1}^m x_i \leq \alpha\Big\}, \qquad B := \Big\{x \in \mathbb{R}^m :\ x_i \geq 0 \text{ for all } i,\ \sum_{i=1}^m x_i \leq \beta\Big\},$$

so that $a_t \in A$ and $b_t \in B$ for all t. Thus, at each time t, Player 1 (resp. Player 2) allocates a nonnegative m-vector of drift intensities across the battlefields, with total instantaneous expense bounded by α (resp. β).

As is standard in game theory, we seek to characterize a Nash equilibrium for this stochastic differential game. For $(t, x_1, x_2) \in [0, T]\times\mathbb{R}^m\times\mathbb{R}^m$, denote by $J(t, x_1, x_2, a, b)$ the payoff obtained

when Player 1 uses a and Player 2 uses b. The upper and lower values of the game are then defined by

$$V^+(t,x_1,x_2) := \inf_{b\in\mathcal{B}} \sup_{a\in\mathcal{A}} J(t,x_1,x_2,a,b), \qquad V^-(t,x_1,x_2) := \sup_{a\in\mathcal{A}} \inf_{b\in\mathcal{B}} J(t,x_1,x_2,a,b).$$

We show below that $V^+ = V^-$, so that the game admits a value, and that it is attained at a pair of strategies $(a^\star, b^\star)$ forming a Nash equilibrium.

Summary of the approach. The symmetric case, where both players have the same budget, behaves qualitatively differently from the asymmetric case $\alpha \neq \beta$. In the symmetric setting, we use a dynamic programming argument and analyze the associated Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation. By symmetry, we postulate that the value function depends only on the difference of the two players' state processes. Under this reduction, the HJBI equation collapses to the linear heat equation, whose regularity is classical. A verification argument then allows to carefully confirm that the value indeed depends only on the difference.

In the asymmetric case, the HJBI equation becomes semilinear. In this regime, it is natural that the value and equilibrium strategies admit a probabilistic characterization in terms of a Backward Stochastic Differential Equation (BSDE). This representation then lends itself to numerical approximation : we provide an algorithm based on a neural network approximation of the BSDE solution, which yields both the value and the equilibrium controls.

In both regimes, both players concentrate their entire budget on the same battlefield.

Main contributions. In this chapter, we extend the classical Colonel Blotto framework in several directions. First, we introduce a time–dynamic version of Blotto games, in which both players observe the state processes and adjust their allocations continuously in time. In addition, we consider a stochastic formulation, where the players control the drift of a stochastic differential equation.

Second, we work under general assumptions on the gain functional : the regularity and structural conditions imposed on the battlefield rewards are kept mild (see Assumption 2.2.1).

Third, we allow for interactions between battlefields. In the standard Blotto model, the total reward is additive across battlefields, which implies that only marginal distributions matter at equilibrium. Our formulation accommodates non-additive criteria, thereby covering a broader class of applications. In particular, the framework encompasses cyber–security models where the reward depends on the algebraic connectivity of a graph (see Section 2.4.4).

Within this general setting, we characterize Nash equilibria : in the symmetric case via the solution of a heat equation, and in the asymmetric case via the solution of a BSDE.

In both regimes, the equilibrium strategies are of bang–bang type : both players allocate all their resources to a single battlefield at each time. Notably, implementing such strategies does not require observing which battlefield the opponent is currently investing in. Note that in standard Colonel Blotto games, no pure Nash equilibrium exists, and one must resort to mixed strategies to obtain equilibrium. In contrast, in our time–dynamic stochastic formulation, no relaxation is needed : pure strategies exist and yield a Nash equilibrium.

Structure of the chapter. Section 2.2 formally introduces the game in weak formulation, together with the admissible controls and the assumptions on the reward. Section 2.3 is devoted to the symmetric case of equal budgets, where we solve the game in closed form. In Section 2.4

we examine several explicit examples, including spectral games on graphs. Section 2.5 treats the asymmetric regime : we show that the game admits a value and we characterize it via a BSDE. Finally, we describe a numerical scheme for this BSDE and provide some illustrations.

2.2 Model and Assumptions

The stochastic basis. Throughout the chapter, we fix a finite time horizon $T > 0$ and a number $m \in \mathbb{N}^\star$ of battlefields. Let $\Omega = \mathcal{C}([0,T], \mathbb{R}^m \times \mathbb{R}^m)$ be the canonical space of continuous trajectories defined on $[0,T]$ with values in $\mathbb{R}^m \times \mathbb{R}^m$ and $X = (X^1, X^2)$ let be the canonical process on Ω, i.e. $(X_t^1, X_t^2)(\omega) = (\omega_1(t), \omega_2(t))$, for $t \in [0,T]$, where $\omega = (\omega_1, \omega_2) \in \Omega$. Ω is equipped with the topology of uniform convergence, so that it is a complete separable metric space. Let $\mathcal{F}$ be the Borel σ-algebra on Ω. Let P be the Wiener measure on $(\Omega, \mathbb{F})$. We denote $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ the filtration generated by X and augmented by the P-null sets.

The admissible controls. We define, for $\lambda > 0$, $A_\lambda := \{x \in \mathbb{R}^m, \forall i = 1, ..., m,\ x_i \ge 0$ and $\sum_{i=1}^m x_i \le \lambda\}$. Let $A := A_\alpha$ and $B := A_\beta$ where α and β are positive constants representing respectively the budgets of player 1 and player 2, fixed for the rest of the section. Let $\mathcal{A}$ (resp. $\mathcal{B}$) be the set of A-valued (resp. B-valued) $\mathbb{F}$-progressively measurable processes.

We are going to introduce a game in weak formulation, in the sense that the controls of the two players are exerted via a family of probability measures. More precisely, for $a \in \mathcal{A}$ and $b \in \mathcal{B}$, we define the probability measure $P^{a,b}$ via a change of measure, with the following Radon–Nikodym derivative :

$$\frac{dP^{a,b}}{dP} := \exp\left(\int_0^T a_s \, dX_s^1 + b_s \, dX_s^2 - \frac{1}{2}\int_0^T (|a_s|^2 + |b_s|^2)\, ds\right).$$

In particular, the process $\left(W_t^{1,a},\, W_t^{2,b}\right)$ defined by

$$\left(W_t^{1,a},\, W_t^{2,b}\right) := \left(X_t^1 - \int_0^t a_s \, ds,\, X_t^2 - \int_0^t b_s \, ds\right), \quad 0 \le t \le T,$$

is a $P^{a,b}$-Brownian motion. Notice that under $P^{a,b}$, the canonical process (X_1, X_2) has the law of a Brownian motion with controlled drift, that is, it satisfies

$$\begin{aligned} X_s^1 &= \int_0^s a_u \, du + W_s^{1,a}, \\ X_s^2 &= \int_0^s b_u \, du + W_s^{1,b}. \end{aligned}$$

We now introduce the criterion J of the game. As usual, we fix the initial value of the controlled process at some time $t \ge 0$. Hence, for $t \le s \le T$ and $(x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^m$, we define

$$\begin{aligned} X_s^{1,t,x_1} &= x_1 + X_{s-t}^1, \\ X_s^{2,t,x_2} &= x_2 + X_{s-t}^2. \end{aligned} \tag{2.1}$$

The criterion. We consider a function $g : \mathbb{R}^m \to \mathbb{R}$, which only depends on the difference between the final values of the processes X^1 and X^2. In this sense, the class of games we consider are so-called *ranking games*. The function g represents the gain of Player 1 for a given pair of terminal states.

Throughout the chapter, g will satisfy the following assumptions :

Assumption 2.2.1.

(i) g has at most exponential growth, i.e. there exist positive constants C_1 and C_2 such that $|g(x)| \leq C_1 e^{C_2|x|}$ for all x in $\mathbb{R}^m$, where we slighly abuse notation and use $|.|$ to denote both a norm on $\mathbb{R}^m$ and the absolute value on $\mathbb{R}$.

(ii) The set of discontinuity points of g has zero Lebesgue measure.

In Section 2.4, we provide several examples of gain functions satisfying Assumption 2.2.1.

Definition 2.2.2. *We define the game criterion as :*

$$J(t, x_1, x_2, a, b) := \mathbb{E}^{P^{a,b}}\left[g\left(X_T^{1,t,x_1} - X_T^{2,t,x_2}\right)\right]$$

where $(X_u^{1,t,x_1})_{t\leq u\leq T}$ and $(X_u^{2,t,x_2})_{t\leq u\leq T}$ are given by (2.1).

In this zero-sum game, Player 1 aims to maximize J, while Player 2 aims to minimize it. In game theory, a *Nash equilibrium* is the standard concept used to describe the solution of a game. It is a strategy profile, i.e. a control for each player, where neither player benefits from unilaterally deviating.

Definition 2.2.3. *For a given game state $(x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^m$ at time $t \in [0, T]$, a pair of strategies $(a_0, b_0) \in (\mathcal{A}, \mathcal{B})$ is called a Nash equilibrium if :*

— $J(t, x_1, x_2, a_0, b_0) \leq J(t, x_1, x_2, a_0, b)$ *for all* $b \in \mathcal{B}$.

— $J(t, x_1, x_2, a, b_0) \leq J(t, x_1, x_2, a_0, b_0)$ *for all* $a \in \mathcal{A}$.

In the next section, we characterize a Nash equilibrium in the symmetric case, i.e. when $\alpha = \beta$.

2.3 The symmetric case

Let us consider the game where both players have the same budget. Without loss of generality, we set $\alpha = \beta = 1$ for the rest of this section. We have in particular $\mathcal{A} = \mathcal{B}$ and the controls are valued in $A := A_1$.

A Nash equilibrium in this game can be obtained by solving two optimization problems, depending on who plays first. One scenario involves Player 1 maximizing the criterion given Player 2's best action, while the other involves Player 2 minimizing the criterion given Player 1's best action. For a given game state (x_1, x_2) and time t :

$$\begin{aligned} &\sup_{a\in\mathcal{A}} \inf_{b\in\mathcal{A}} J(t, x_1, x_2, a, b) \text{ when Player 2 plays first,} \\ &\inf_{b\in\mathcal{A}} \sup_{a\in\mathcal{A}} J(t, x_1, x_2, a, b) \text{ when Player 1 plays first.} \end{aligned} \tag{2.1}$$

In order to solve these optimization problems, we use the dynamic programming approach, which consists in analyzing the upper and lower Isaacs equations.

Definition 2.3.1. *Let w be the function defined on $[0,T)\times\mathbb{R}^m$ by :*

$$w(t,x) := \frac{1}{(4\pi(T-t))^{\frac{m}{2}}}\int_{\mathbb{R}^m}\exp\left[-\frac{|x-z|^2}{4(T-t)}\right]g(z)dz \tag{2.2}$$

Note that $w\in\mathcal{C}^{1,2}([0,T)\times\mathbb{R}^m)$ and thet w solves the following backward *heat equation*, with terminal condition given by g :

$$\begin{aligned}\partial_t w+\Delta w=0 \text{ in } [0,T)\times\mathbb{R}^m\\ w(T,x)=g(x) \text{ on } \mathbb{R}^m\end{aligned} \tag{2.3}$$

Next, we define a pair of feedback controls depending only on the difference of the players' state processes on each battlefield. Theorem 2.3.6 below shows that this pair of strategies constitutes a Nash equilibrium for the game.

Definition 2.3.2. *For $(t,x_1,x_2)\in[0,T]\times\mathbb{R}^m\times\mathbb{R}^m$, we define*

$$\tilde{a}(t,x_1,x_2) := \begin{cases} e_i \text{ if } \max\limits_{j=1,\dots,m}\partial_j w=\partial_i w>0\\ 0_{\mathbb{R}^m} \text{ otherwise}\end{cases} \qquad \tilde{b}(t,x_1,x_2) := \begin{cases} e_i \text{ if } \min\limits_{j=1,\dots,m}-\partial_j w=\partial_i w<0\\ 0_{\mathbb{R}^m} \text{ otherwise,}\end{cases}$$

where, for each $i=1,\dots,m$, $e_i\in\mathbb{R}^m$ denotes the i-th canonical basis vector, and $\partial_i w$ stands for the partial derivative of w with respect to the i-th spatial coordinate, evaluated at $(t,x_1-x_2)\in[0,T]\times\mathbb{R}^m$.

Definition 2.3.3. *For a given game state $(x_1,x_2)\in\mathbb{R}^m\times\mathbb{R}^m$ at time $t\in[0,T]$, we define the processes $(a^\star)_{t\le s\le T} := \left(\tilde{a}\left(s,X_s^{1,t,x_1},X_s^{2,t,x_2}\right)\right)_{t\le s\le T}$ and $(b^\star)_{t\le s\le T} := \left(\tilde{b}\left(s,X_s^{1,t,x_1},X_s^{2,t,x_2}\right)\right)_{t\le s\le T}$*

The processes $a^\star$ and $b^\star$ implement strategies in which each player, at every time, identifies the battlefield with the largest marginal impact on the expected terminal payoff, and allocates her entire budget to that battlefield.

To prove that the pair $(a^\star,b^\star)$ forms a Nash equilibrium, we will actually establish a stronger result : $(a^\star,b^\star)$ is a saddle point for the game. This is the content of Theorem 2.3.6 below.

Definition 2.3.4. *A pair of strategies (a_0,b_0) is a saddle point of the game if :*

$$J(t,x_1,x_2,a_0,b_0)=\sup_{a\in\mathcal{A}}J(t,x_1,x_2,a,b_0)=\inf_{b\in\mathcal{A}}J(t,x_1,x_2,a_0,b)$$

for all $(t,x_1,x_2)\in[0,T]\times\mathbb{R}^m\times\mathbb{R}^m$. In particular, a_0 and b_0 are mutual best responses and form a Nash equilibrium.

For $v:[0,T]\times\mathbb{R}^m\times\mathbb{R}^m\to\mathbb{R}$, we write

$$\begin{aligned}D_{x_1}v\big(t,x_1,x_2\big)&=\big(\partial_{x_1(i)}v(t,x_1,x_2)\big)_{1\le i\le m} \quad \text{and similarly}\\ D_{x_2}v\big(t,x_1,x_2\big)&=\big(\partial_{x_2(i)}v(t,x_1,x_2))\big)_{1\le i\le m}.\end{aligned}$$

When there is no confusion, we simply write $D_{x_1}v$ or $D_{x_2}v$. In the case where $v(t,x_1,x_2)=w(t,x_1-x_2)$ only depends on x_1-x_2, we will slightly abuse notation and write

$$\begin{aligned}D_{x_1}w=D_{x_1}w(t,x_1-x_2)=Dw(t,x_{1x_2}) \quad \text{and}\\ D_{x_2}w=D_{x_2}w(t,x_1-x_2)=-Dw(t,x_1-x_2),\end{aligned}$$

where Dw denotes the usual gradient of w with respect to the space variable.

The Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation associated with the game reads

$$\begin{gathered}\partial_t u + \frac{1}{2}\Delta u + \sup_{a\in A} a.D_{x_1}u + \inf_{b\in A} b.D_{x_2}u = 0 \text{ in } [0,T)\times\mathbb{R}^m\times\mathbb{R}^m \\ u(T,x_1,x_2) = g(x_1-x_2) \text{ on } \mathbb{R}^m\times\mathbb{R}^m.\end{gathered}$$

It simplifies to :

$$\begin{gathered}\partial_t u + \frac{1}{2}\Delta u + \left(\max_{i=1,\dots,m}\partial_{x_1(i)}u\right)^+ + \left(\min_{i=1,\dots,m}\partial_{x_2(i)}u\right)^- = 0 \text{ in } [0,T)\times\mathbb{R}^m\times\mathbb{R}^m \\ u(T,x_1,x_2) = g(x_1-x_2) \text{ on } \mathbb{R}^m\times\mathbb{R}^m,\end{gathered} \tag{2.4}$$

where $(s)^+ = \max(s,0)$ and $(s)^- = \min(s,0)$ for $s\in\mathbb{R}$.

Definition 2.3.5. *We define the upper (resp. lower) value function V^+ (resp. V^-) as :*

$$V^+(t,x_1,x_2) = \inf_{b\in\mathcal{A}}\sup_{a\in\mathcal{A}} J(t,x_1,x_2,a,b) \tag{2.5}$$

$$V^-(t,x_1,x_2) = \sup_{a\in\mathcal{A}}\inf_{b\in\mathcal{A}} J(t,x_1,x_2,a,b) \tag{2.6}$$

Theorem 2.3.6. *The upper and lower value functions coincide and we have, $\forall(t,x_1,x_2)\in[0,T]\times\mathbb{R}^m\times\mathbb{R}^m$,*

$$V^+(t,x_1,x_2) = V^-(t,x_1,x_2) = J(t,x_1,x_2,a^\star,b^\star) = w(t,x_1-x_2),$$

where w is defined in (2.2)*. Hence the value function $V := V^+ = V^-$ is $\mathcal{C}^{1,2}$, solves the HJBI equation* (2.4) *and the pair of strategies $(a^\star,b^\star)$ is a Nash equilibrium of the game.*

Proof. The proof of this theorem is given by Lemmas 2.3.8 and 2.3.9 below. □

The next lemma shows that w, which depends on (x_1,x_2) only through the difference x_1-x_2, solves the HJBI equation (2.4). Lemma 2.3.8 then provides the key step ensuring that the game admits a value, and that this value indeed depends only on x_1-x_2.

Lemma 2.3.7. *We have for all $(t,x_1,x_2)\in[0,T)\times\mathbb{R}^m\times\mathbb{R}^m$ and $\gamma = x_1-x_2$:*

$$\begin{gathered}\partial_t w(t,\gamma) + \Delta w(t,\gamma) + \sup_{a\in A} a.D_x w(t,\gamma) + \inf_{b\in A} b.D_y w(t,\gamma) = 0 \\ \partial_t w(t,\gamma) + \Delta w(t,\gamma) + a^\star(t,x_1,x_2).D_x w(t,\gamma) + \inf_{b\in A} b.D_y w(t,\gamma) = 0 \\ \partial_t w(t,\gamma) + \Delta w(t,\gamma) + \sup_{a\in A} a.D_x w(t,\gamma) + b^\star(t,x_1,x_2).D_y w(t,\gamma) = 0\end{gathered} \tag{2.7}$$

Proof. We first notice that the three equations are equivalent since for $(t,x_1,x_2)\in[0,T)\times\mathbb{R}^m\times\mathbb{R}^m$, we have

$$\sup_{a\in A} a.D_{x_1}w(t,x_1-x_2) = \left(\max_{i=1,\dots,m}\partial_i w(t,\gamma)\right)^+$$

Where $\gamma := x_1-x_2$, and

$$\inf_{b\in A} b.D_{x_2}w(t,\gamma) = \left(\min_{i=1,\dots,m}\partial_{x_2,i}w(t,\gamma)\right)^- = \left(\min_{i=1,\dots,m} -\partial_i w(t,\gamma)\right)^-$$

Finally, for $(t, x_1, x_2) \in [0,T) \times \mathbb{R}^m \times \mathbb{R}^m$ the first equation becomes :

$$\partial_t w(t,\gamma) + \Delta w(t,\gamma) + \left(\max_{i=1,...,m} \partial_i w(t,\gamma)\right)^+ + \left(\min_{i=1,...,m} -\partial_i w(t,\gamma)\right)^- = \partial_t w(t,\gamma) + \Delta w(t,\gamma) = 0$$

Using the equation (2.3) and noticing that $\forall q_1, ..., q_n \in \mathbb{R};\ \min(0, q_1, ..., q_n) = -\max(0, -q_1, ..., -q_n)$. □

Lemma 2.3.8. *The pair $(a^\star, b^\star)$ is a saddle point, and we have :*

$$w(t, x_1 - x_2) = J(t, x_1, x_2, a^\star, b^\star) = \sup_{a \in \mathcal{A}} J(t, x_1, x_2, a, b^\star) = \inf_{b \in \mathcal{A}} J(t, x_1, x_2, a^\star, b)$$

for all $(t, x_1, x_2) \in [0,T] \times \mathbb{R}^m \times \mathbb{R}^m$.

The proof of Lemma 2.3.8, which uses a standard but technical verification argument, is deferred to the Appendix.

Lemma 2.3.9. *The upper and lower value functions coincide and we have :*

$$\forall (t, x_1, x_2) \in [0,T] \times \mathbb{R}^m \times \mathbb{R}^m; V^+(t, x_1, x_2) = V^-(t, x_1, x_2) = w(t, x_1 - x_2)$$

The value function $V := V^+ = V^-$ solves the HJBI equation 2.4.

Proof. Lemma 2.3.8 gives for all $(t, x_1, x_2) \in [0,T] \times \mathbb{R}^m \times \mathbb{R}^m$:

$$\begin{aligned} V^+(t, x_1, x_2) &\leq \sup_{a \in \mathcal{A}} J(t, x_1, x_2, a, b^\star) = w(t, x_1 - x_2) \\ &= \inf_{b \in \mathcal{A}} J(t, x_1, x_2, a^\star, b) \leq V^-(t, x_1, x_2) \leq V^+(t, x_1, x_2). \end{aligned}$$

Hence $V^+ = V^-$, and by Lemma 2.3.7, the value function V solves the HJBI equation (2.4). □

2.4 Particular cases in the symmetric setting

In this section, we exploit that, in the symmetric case, the value function coincides with the smooth solution $w \in \mathcal{C}^{1,2}$ of the backward heat equation (2.3) and depends only on the difference $x_1 - x_2$. This observation allows us to compute w—and the associated bang–bang equilibrium controls—explicitly for several representative classes of payoffs.

2.4.1 Fixed payoff for each battlefield

With the same notations, we consider the problem above in the case where $g(x) = \sum_{i=1}^m \alpha_i \mathbf{1}_{\mathbb{R}^+}(x_i)$ for all $x \in \mathbb{R}^m$, where $(\alpha_i)_{i=1,...,m}$ are real numbers. This is the payoff function of the standard colonel Blotto games. The function g satisfies Assumption 2.2.1 : g is bounded so it has at most exponential growth, and its set of discontinuity is the union of the hyperplanes $\bigcup_{i=1}^m \{x \in \mathbb{R}^m;\ x_i = 0\}$ which is a null set with regards to the Lebesgue measure.
In this particular case, we can compute w explicitly using the Feynman-Kac formula and computing the expectation $\mathbb{E}\left[g\left(x + \sqrt{2}B_{T-t}\right)\right]$, where B is an m-dimensional Brownian motion. We have for all $(t, x) \in [0,T) \times \mathbb{R}^m$:

$$w(t,x) = \sum_{i=1}^m \alpha_i \Phi\left(\sqrt{\frac{1}{2(T-t)}} x_i\right)$$

where Φ is the cumulative distribution function of the standard normal distribution. This means that for all $(t,x) \in [0,T) \times \mathbb{R}^m$:

$$\partial_i w(t,x) = \frac{\alpha_i}{\sqrt{4\pi(T-t)}} \exp\left[-\frac{x_i^2}{4(T-t)}\right]$$

The case where all the (α_i) are equal, meaning all battlefields have the same weight, the strategies where both players focus all their budget on the closest battle is a Nash equilibrium. Otherwise, when $t << T$ the α_i dominate when the differences between players' processes aren't too big, and when $t \to T$ the α_i are less important and the differences between the processes play a bigger role in choosing which battlefield to focus on.
Looking at the point where the players would transfer their budget to another battlefield :

$$\partial_i w(t,x) \leq \partial_j w(t,x) \iff x_i^2 - x_j^2 \geq 4(T-t)\log\left(\frac{\alpha_i}{\alpha_j}\right)$$

So in the case where all battlefields share the same weight, this translates to both players focusing the entirety of their budgets on the battlefields where the difference between the state processes in the smallest in absolute value.

2.4.2 Elections inspired payoff

Let us consider the case where for all $x \in \mathbb{R}^m$:

$$g(x) = \mathbf{1}_{\mathbb{R}_+}\left(\sum_{i=1}^m \mathbf{1}_{\mathbb{R}_+}(x_i) - \frac{m}{2}\right)$$

This would correspond to the US presidential elections model for example, where battlefields would correspond to states.
This function g satisfies Assumption 2.2.1 : g is bounded so it has at most exponential growth, and its set of discontinuity is $\left\{x \in \mathbb{R}^m;\ \sum_{i=1}^m \mathbf{1}_{\mathbb{R}_+}(x_i) = \frac{m}{2}\right\}$ which is a null set with regards to the Lebesgue measure.
We can explicitly characterize the w function here as well, we have for all $(t,x) \in [0,T) \times \mathbb{R}^m$:

$$w(t,x) = \sum_{j=\lceil \frac{m}{2} \rceil}^m \left(\sum_{p_1<p_2<\ldots<p_j \in [\![1,m]\!]} \prod_{k=1}^j \Phi\left(\sqrt{\frac{1}{2(T-t)}} x_{p_k}\right)\right)$$

And the partial derivatives give for all $i \in [\![1,m]\!]$ and $(t,x) \in [0,T) \times \mathbb{R}^m$:

$$\partial_i w(t,x) = \frac{\exp\left[-\frac{x_i^2}{4(T-t)}\right]}{\sqrt{4\pi(T-t)}} \sum_{j=\lceil \frac{m}{2} \rceil}^m \left(\sum_{p_1<p_2<\ldots<p_{j-1} \in [\![1,m]\!]} \prod_{k=1}^{j-1} \Phi\left(\sqrt{\frac{1}{2(T-t)}} x_{p_k}\right)\right)$$

We can characterize the points where players' strategies would flip between battlefields, but there is less of an intuitive interpretation. We have $\partial_i w(t,x) = \partial_l w(t,x)$ is equivalent to :

$$\frac{\Phi}{\phi}\left(\sqrt{\frac{1}{2(T-t)}} x_l\right) \frac{\phi}{\Phi}\left(\sqrt{\frac{1}{2(T-t)}} x_i\right) K_1^{j,l}(x) + \frac{\phi}{\Phi}\left(\sqrt{\frac{1}{2(T-t)}} x_i\right) K_2^{j,l}(x) = K_1^{j,l}(x) + \frac{1}{\Phi\left(\sqrt{\frac{1}{2(T-t)}} x_i\right)} K_2^{j,l}(x)$$

for all $(t,x) \in [0,T) \times \mathbb{R}^m$ and $i,j \in [\![1,m]\!]$, where ϕ is the standard normal distribution's density function, and :

$$\begin{cases} K_1^{j,l}(x) = \sum_{j=\lceil \frac{m}{2} \rceil}^{m} \left(\sum_{p_1<p_2<\ldots<p_{j-2} \in [\![1,m]\!] \setminus \{i,j\}} \prod_{k=1}^{j-2} \Phi\left(\sqrt{\frac{1}{2(T-t)}} x_{p_k} \right) \right) \\ K_2^{j,l}(x) = \sum_{j=\lceil \frac{m}{2} \rceil}^{m} \left(\sum_{p_1<p_2<\ldots<p_{j-1} \in [\![1,m]\!] \setminus \{i,j\}} \prod_{k=1}^{j-1} \Phi\left(\sqrt{\frac{1}{2(T-t)}} x_{p_k} \right) \right) \end{cases}$$

2.4.3 Winning all the battlefields

We now consider the case where Player 1 receives a positive payoff only if she wins all battlefields. In this setting the gain function takes the form

$$g(t,x) = C \prod_{i=1}^{m} \mathbf{1}_{\mathbb{R}_+}(x_i), \qquad (t,x) \in [0,T) \times \mathbb{R}^m,$$

where $C > 0$ is a constant. We have in this case for all $(t,x) \in [0,T) \times \mathbb{R}^m$ and $i \in [\![1,m]\!]$:

$$w(t,x) = C \prod_{j=1}^{m} \Phi\left(\sqrt{\frac{1}{2(T-t)}} x_j \right)$$

and

$$\partial_i w(t,x) = C \frac{\exp\left[-\frac{x_i^2}{4(T-t)} \right]}{\sqrt{4\pi(T-t)}} \prod_{j \in [\![1,m]\!] \setminus \{i\}} \Phi\left(\sqrt{\frac{1}{2(T-t)}} x_j \right) = \sqrt{\frac{1}{2(T-t)}} C \frac{\phi\left(\sqrt{\frac{1}{2(T-t)}} x_i \right)}{\Phi\left(\sqrt{\frac{1}{2(T-t)}} x_i \right)} w(t,x)$$

This gives :

$$\partial_i w(t,x) = \partial_j w(t,x) \iff \frac{\phi\left(\sqrt{\frac{1}{2(T-t)}} x_i \right)}{\Phi\left(\sqrt{\frac{1}{2(T-t)}} x_i \right)} = \frac{\phi\left(\sqrt{\frac{1}{2(T-t)}} x_j \right)}{\Phi\left(\sqrt{\frac{1}{2(T-t)}} x_j \right)}$$

The standard normal distribution's hazard function $\frac{\phi}{\Phi}$ is strictly decreasing (and continuous). We get :

$$\partial_i w(t,x) < \partial_j w(t,x) \iff x_i > x_j$$

this means that the changes in battlefield focus occur whenever the components of z switch order. In particular, at equilibrium both players concentrate their resources on the battlefield where Player 1 is currently most disadvantaged (or least ahead). This is consistent with the structure of the payoff : since Player 1 receives a reward only if she wins every battlefield, the optimal response is to defend first the weakest link, i.e., the battlefield most likely to be lost.

2.4.4 Spectral games on graphs

In this subsection, we illustrate the framework with examples arising in cyber-security. We consider competitive situations on networks represented by graphs, where the objective functions depend on the spectrum of the associated Laplacian matrices.
We model a network as an undirected graph $G = (V,E)$, where V denotes the set of vertices and E the set of edges, with respective cardinalities $|V| = n \geq 2$ and $|E| = m \geq 1$. As usual, E is a set of unordered pairs of vertices. For convenience, we index the edges as $E = \{1,\ldots,m\}$, and we assume throughout that the graph has neither loops nor multiple edges.

For $(x_1, x_2) \in \mathbb{R}^m \times \mathbb{R}^m$, we define the weighted graph $G_{x_1-x_2} = (V, E, w)$, where the weight vector $w = (w_j)_{1\le j\le m}$ is given by

$$w_j := \begin{cases} 1, & \text{if } x_1(j) - x_2(j) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Equivalently, $G_{x_1-x_2}$ is the *spanning subgraph* of G with edge set $\widetilde{E} \subset E$ defined by

$$\widetilde{E} = \{\, e_j \in E \ :\ x_1(j) - x_2(j) > 0 \,\}.$$

Let L denote the Laplacian matrix associated with $G_{x_1-x_2}$, defined as

$$L = D - A,$$

where A is the adjacency matrix and D the degree matrix, i.e., the diagonal matrix with entries D_{ii} equal to the degree of vertex i. The matrix L is real, symmetric, and positive semidefinite, and always admits 0 as an eigenvalue.
Let $\lambda_1 = 0 \le \lambda_2 \le \cdots \le \lambda_n$ denote the eigenvalues of L. We consider *spectral objective functions*, namely functions ψ that depend on the Laplacian spectrum and can be written in the form $\psi(\lambda_1, \dots, \lambda_n) = \phi(w)$.
Given a graph $G = (V, E)$, we define the payoff function $g : \mathbb{R}^m \to \mathbb{R}$ by $g(v) := \psi(G_v)$. In words, each edge of G corresponds to a battlefield. If Player 1 (who aims to maximize her spectral criterion) wins battlefield j, then the corresponding edge e_j remains in the graph ; otherwise, it is removed.
Player 1 therefore allocates her control dynamically across the edges in order to maximize

$$\mathbb{E}^{P^{a,b}}\left[\psi\Big(G_{X_T^{1,t,x_1} - X_T^{2,t,x_2}}\Big)\right],$$

while Player 2 seeks to minimize this same quantity.

Theorem 2.4.1. *The value function of the spectral game is given explicitly by*

$$V(t, x_1, x_2) = \sum_{\widetilde{G}\in\mathcal{S}} \psi(\widetilde{G})\, \Phi_{\widetilde{G}}(T - t, x_1 - x_2), \tag{2.1}$$

where $\mathcal{S}$ denotes the set of spanning subgraphs of G,

$$\Phi_{\widetilde{G}}(s, v) := \prod_{j\in\widetilde{E}} \Phi\Big(\frac{v_j}{\sqrt{2s}}\Big) \prod_{j\in\widetilde{E}^c} \overline{\Phi}\Big(\frac{v_j}{\sqrt{2s}}\Big), \quad s > 0,\ v \in \mathbb{R}^m, \tag{2.2}$$

and where we recall that Φ denotes the cumulative distribution function of a standard Gaussian distribution on $\mathbb{R}$ and $\overline{\Phi} := 1 - \Phi$.

Moreover, for $w(t, x_1 - x_2) := V(t, x_1, x_2)$, the spatial derivatives are given by

$$\partial_i w(t, x) = \sum_{\widetilde{G}\in\mathcal{S}} \frac{\psi(\widetilde{G})\,(-1)^{\mathbf{1}_{\{i\in\widetilde{E}^c\}}}}{\sqrt{2(T-t)}}\, \phi\Big(\frac{x_i}{2(T-t)}\Big) \prod_{j\in\widetilde{E}\setminus\{i\}} \Phi\Big(\frac{x_j}{\sqrt{2(T-t)}}\Big) \prod_{j\in\widetilde{E}^c\setminus\{i\}} \overline{\Phi}\Big(\frac{x_j}{\sqrt{2(T-t)}}\Big).$$

Connectivity game on a network

Let us now consider a competition between two agents on a network : the first agent seeks to protect the network, aiming to keep it as connected as possible, while the second agent tries to disconnect it. We begin by recalling a standard measure of graph connectedness.

Definition 2.4.2. *The second smallest eigenvalue $\lambda_2 \geq 0$ of the Laplacian matrix L is called the algebraic connectivity. When needed, we write $\lambda_2(G)$ to emphasize its dependence on the underlying graph G.*

This quantity captures several aspects of a graph's connectedness. In particular, $\lambda_2 = 0$ if and only if the graph is disconnected. Furthermore, adding an edge to a connected graph cannot decrease the algebraic connectivity, and therefore this quantity is maximal for the complete graph. We refer to [75] and the references therein for more details on the spectral properties of Laplacian matrices.

Remark 2.4.3. *One can verify directly from* (2.1) *that*

$$V(T, x_1, x_2) = \lambda_2 \left(G_{x_1 - x_2}\right).$$

Indeed, taking $t = T$ in (2.1) *yields*

$$V(T, x_1, x_2) = \sum_{\widetilde{G} \in \mathcal{S}} \psi(\widetilde{G})\, \Phi_{\widetilde{G}}(0, x_1 - x_2).$$

By construction, the function $\Phi_{\widetilde{G}}(s, v)$ in (2.1) *acts as the probability weight that the edge configuration $\widetilde{G}$ arises at time horizon $T - t = s$. In the limit $s \to 0$, this weight converges to an indicator function selecting precisely the subgraph associated with the realized sign pattern of $x_1 - x_2$:*

$$\lim_{s \to 0} \Phi_{\widetilde{G}}(s, v) = \mathbf{1}_{\{\widetilde{G} = G_v\}}.$$

Substituting this limit into the previous expression gives

$$V(T, x_1, x_2) = \sum_{\widetilde{G} \in \mathcal{S}} \psi(\widetilde{G})\, \mathbf{1}_{\{\widetilde{G} = G_{x_1 - x_2}\}} = \psi(G_{x_1 - x_2}).$$

Recalling that, in the present setting, $\psi(G)$ is chosen as the algebraic connectivity $\lambda_2(G)$ of the graph G, we obtain

$$V(T, x_1, x_2) = \lambda_2(G_{x_1 - x_2}),$$

as claimed.

Corollary 2.4.4. *In the algebraic connectivity game, the value function is given by*

$$V(t, x_1, x_2) = \sum_{\widetilde{G} \in \mathcal{S}^\rfloor} \lambda_2(\widetilde{G})\, \Phi_{\widetilde{G}}(T - t, x_1 - x_2),$$

where $\mathcal{S}^\rfloor$ is the set of connected spanning subgraphs of G, and $\Phi_{\widetilde{G}}$ is defined as in (2.1).

The quantity $V(t, x_1, x_2)$ thus represents the expected algebraic connectivity of the network at time T, given the current difference $x_1 - x_2$ in the players' state processes and under the equilibrium strategies. Since $\lambda_2(G)$ increases when edges are added and vanishes when the graph

becomes disconnected, this value can be seen as a measure of the network's expected robustness at equilibrium.

From a game-theoretic perspective, Player 1 concentrates her control on the edges that contribute the most to increasing $\lambda_2(G_{x_1-x_2})$, i.e., those that connect otherwise weakly linked components of the network. Conversely, Player 2 focuses on edges whose removal most reduces λ_2, typically the "bottlenecks" or critical links in the topology.

In the symmetric case, where both players have equal budgets, the resulting equilibrium corresponds to a dynamic tug-of-war on these structurally important edges. The equilibrium strategies, derived from the gradient of the value function $w(t, x_1 - x_2)$, therefore encode an adaptive defense–attack pattern over the network : each player continuously reallocates effort toward the edges with the greatest instantaneous impact on the network's algebraic connectivity.

2.5 The asymmetric case

In this section, we consider the case where players don't have the same budget i.e. $\alpha \neq \beta$. We can still characterize a Nash equilibrium, this time using Backward Stochastic Differential Equations (BSDEs), following the approach of [36].

2.5.1 Characterization of a Nash Equilibrium

Definition 2.5.1. *We introduce the Hamiltonian function $H : \mathbb{R}^m \times A \times B \to \mathbb{R}$ as :*

$$H(z, a, b) := (a - b).z$$

It is clear that the following condition is satisfied for all $z \in \mathbb{R}^m$:

$$\inf_{b\in B} \sup_{a\in A} H(z, a, b) = \sup_{a\in A} \inf_{b\in B} H(z, a, b)$$

Definition 2.5.2. *We introduce the following controls for all $z \in \mathbb{R}^m$:*

$$\tilde{a}(z) = \begin{cases} \alpha e_i \text{ if } \max\limits_{j=1,\dots,m} z_j = z_i > 0 \\ 0_{\mathbb{R}^m} \text{ otherwise} \end{cases} \qquad \tilde{b}(z) = \begin{cases} \beta e_i \text{ if } \min\limits_{j=1,\dots,m} - z_j = z_i < 0 \\ 0_{\mathbb{R}^m} \text{ otherwise} \end{cases}$$

where for all $i = 1, ..., m$, $e_i \in \mathbb{R}^m$ is the ith vector of the canonical base of $\mathbb{R}^m$ and z_i is the ith coordinate of the vector z.

We define $\widetilde{\mathcal{A}}$ (respectively $\widetilde{\mathcal{B}}$) as the set of measurable functions from $\mathbb{R}^m$ with values in A (respectively B). In particular, $\tilde{a} \in \mathcal{A}$ and $\tilde{b} \in \mathcal{B}$.

Lemma 2.5.3. *The processes $\tilde{a}$ and $\tilde{b}$ are such that for all $(z, a, b) \in \mathbb{R}^m \times A \times B \to \mathbb{R}$:*

$$H\left(z, a, \tilde{b}(z)\right) \leq H\left(z, \tilde{a}(z), \tilde{b}(z)\right) \leq H\left(z, \tilde{a}(z), b\right)$$

And we have :

$$H\left(z, \tilde{a}(z), \tilde{b}(z)\right) = \inf_{b\in B} \sup_{a\in A} H(z, a, b) = \sup_{a\in A} \inf_{b\in B} H(z, a, b)$$

Proof. We prove for $\tilde{a}$ first. Let $(a,b) \in A \times B$ and $p \in \mathbb{R}^m$, we have :

$$\begin{aligned} H(z,a,b) = (a-b).z &= -b.z + \sum_{i=1}^{m} a_i z_i \\ &\leq -b.z + \sum_{i=1}^{m} a_i \max_{j \in [\![1,m]\!]} z_j \\ &\leq -b.z + \max_{j \in [\![1,m]\!]} z_j \sum_{i=1}^{m} a_i \\ &\leq -b.z + \alpha \max_{j \in [\![1,m]\!]} z_j \\ &\leq \tilde{a}(z).z - b.z \\ &= H\left(z, \tilde{a}(z), b\right) \end{aligned}$$

We get in a similar way that $H\left(z,a,\tilde{b}(z)\right) \leq H(z,a,b)$ which concludes the proof. □

Going forward, in order to characterize a Nash equilibrium, we need to introduce what a BSDE is. Given a terminal random variable ξ and a generator function $f : [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, the solution to a Backward Stochastic Differential Equations (BSDEs) is the pair of adapted processes $(Y_t, Z_t)_{0 \leq t \leq T}$ that satisfy :

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dM_s$$

Where M_t is an m-dimensional martingale. In the rest of this chapter, we will take $M = X^1 - X^2$ and specify initial conditions (see for example lemma 2.5.4), which would be a particular P-martingale having the law of a Brownian multiplied by $\sqrt{2}$. The terminal condition is $Y_T = \xi$. The generator f can depend on the current time t, the value of the process $Y_t \in \mathbb{R}$, and the process $Z_t \in \mathbb{R}^m$. The solution (Y_t, Z_t) describes the backward evolution of the system, starting from the terminal state ξ and evolving back to $t = 0$. Under some regularity conditions, namely that the generator function is Lipschitz-continuous and the terminal condition is square integrable, the BSDE has a unique solution (Y^{sol}, Z^{sol}) in $\mathcal{S}^2([0,T)) \times \mathcal{H}^2_m([0,T))$. Where $\mathcal{S}^2([0,T)) := \left\{\phi \in \mathcal{P}_1;\ \mathbb{E}\left[\sup_{s \leq T} |\phi_s|^2\right] < \infty\right\}$ and $\mathcal{H}^2_m([0,T)) := \left\{\phi \in \mathcal{P}_m;\ \mathbb{E}\left[\int_0^T |\phi_s|^2 ds\right] < \infty\right\}$, with $\mathcal{P}_i$ is the set of i-dimensional $\mathbb{F}$-progressively measurable processes for all $i \geq 1$.

Lemma 2.5.4. *Let $(a,b) \in \mathcal{A} \times \mathcal{B}$ and $\left(Y^{a,b}, Z^{a,b}\right)$ be the unique solution of the Backward Stochastic Differential Equation (BSDE) with generator $z \to (a-b).z$ and terminal condition $g\left(X_T^{1,t,x_1} - X_T^{2,t,x_2}\right)$ for $0 \leq t \leq T$ and $(x_1,x_2) \in \mathbb{R}^m \times \mathbb{R}^m$.*
Then $J(t,x_1,x_2,a,b) = \mathbb{E}^{P^{a,b}}\left[g(X_T^{1,t,x_1} - X_T^{2,t,x_2})|\mathcal{F}_t\right] = Y_t^{a,b}$.

Proof. Since H is linear (and thus Lipschitz) in z, the solution $\left(Y^{a,b}, Z^{a,b}\right)$ exists and is unique in $\mathcal{S}^2([0,T)) \times \mathcal{H}^2_m([0,T))$.
Let $t \leq s \leq T$, we have by definition :

$$Y_s^{a,b} = g(X_T^{1,t,x_1} - X_T^{2,t,x_2}) + \int_s^T Z_u^{a,b}.(a-b)\, du - \int_s^T Z_u^{a,b} d\left(X_u^{1,t,x_1} - X_u^{2,t,x_2}\right)$$

Using the earlier defined $\left(W_t^{1,a},\, W_t^{2,b}\right) := \left(X_t^1 - \int_0^t a_s\, ds,\, X_t^2 - \int_0^t b_s\, ds\right)$, we can write

$$Y_s^{a,b} = Y_t^{a,b} + \int_t^s Z_u^{a,b}\left(dW_u^{1,a} - dW_u^{2,b}\right)$$

So $Y^{a,b}$ is a $P^{a,b}$ martingale. This gives us :

$$J(t,x_1,x_2,a,b) = \mathbb{E}^{P^{a,b}}\left[g(X_T^{1,t,x_1} - X_T^{2,t,x_2})|\mathcal{F}_t\right] = \mathbb{E}^{P^{a,b}}\left[Y_T^{a,b}|\mathcal{F}_t\right] = Y_t^{a,b}$$

□

Lemma 2.5.5. *Let* $\left(\tilde{Y},\tilde{Z}\right)$ *be the solution of the BSDE with generator*

$z \to z.\left(\tilde{a}(z) - \tilde{b}(z)\right)$ *and terminal condition* $g\left(X_T^{1,t,x_1} - X_T^{2,t,x_2}\right)$ *for* $(x_1,x_2) \in \mathbb{R}^m \times \mathbb{R}^m$ *;* $0 \le t \le s \le T$

We then define $(a^\star,b^\star) \in \mathcal{A}\times\mathcal{B}$ *as* $(a_s^\star,b_s^\star) := \left(\tilde{a}\left(\tilde{Z}_s\right),\tilde{b}\left(\tilde{Z}_s\right)\right)$
Then $J(t,x_1,x_2,a^\star,b^\star) = \tilde{Y}_t$ *and* $(a^\star,b^\star)$ *is a saddle point of the game.*

Proof. The first result is due to lemma 2.5.4, by noticing that (Y^*,Z^*) solve the same BSDE as $\left(\tilde{Y},\tilde{Z}\right)$.
To prove that $(a^\star,b^\star)$ is a saddle point, we first prove that $a^\star$ is an optimal response to $b^\star$. Let $(x_1,x_2) \in \mathbb{R}^m\times\mathbb{R}^m$, $a \in \mathcal{A}$ and $0 \le t \le s \le T$. Let $(Y^\star,Z^\star)$ (resp. (Y,Z)) be the unique solution of the BSDE with generator $(t,\omega,z) \to z.(a_t^\star(\omega) - b_t^\star(\omega))$ (resp. $z \to z.(a,b^\star)$) and terminal condition $g\left(X_T^{1,t,x_1} - X_T^{2,t,x_2}\right)$.
Using the result of lemma 2.5.3 and the comparison theorem in [65], we have P-almost surely :

$$\forall (t,x_1,x_2) \in [0,T]\times\mathbb{R}^m\times\mathbb{R}^m; Y_t \le Y_t^\star$$

Meaning for t=0 :

$$\forall (x_1,x_2) \in \mathbb{R}^m\times\mathbb{R}^m; J(t,x_1,x_2,a,b^\star) = Y_0 \le Y_0^\star = J(0,x_1,x_2,a^\star,b^\star)$$

We deduce the $b^\star$'s optimality in a similar manner, which proves that $(a^\star,b^\star)$ is a saddle point of the game. □

2.5.2 Numerics

Description of the numerical scheme

In this section, we provide numerical results using neural networks based algorithms to solve BSDEs. The scheme we use is a variation of the *Deep BSDE* scheme originally proposed by E–Han–Jentzen [26].
Let $t_0 = 0 < \cdots < t_N = T$ be a given time discretization. Given a starting point $x \in \mathbb{R}^m$ and a BSDE with generator f and terminal condition function g, we define :

- a learnable scalar parameter $y_0 \in \mathbb{R}$, i.e. a minimizer in the loss optimization given in (2.1) below, used to approximate Y_0,
- a neural network : $\psi_Z(t,x,\theta_Z) : [0,T)\times\mathbb{R}^m\times\mathbb{R}^{d_Z} \to \mathbb{R}^m$, used to approximate Z, when $X_t = x$.

Trajectory simulations We first simulate M trajectories of X using an Euler-Maruyama scheme, depending of simulations of paths of the Brownian motion W, and we denote X_i^l and W_i^l respectively the lth path simulation of X and W taken at time t_i $(0 \le l \le M)$. Let : $\Delta W_i^l := W_{i+1}^l - W_i^l$.

Then starting from $Y_0 := y_0$, we propagate forward the BSDE using the formula

$$Y_{i+1}^l = Y_i^l - f\big(t_i,X_i^l,Y_i^l,\psi_Z(t_i,X_i^l,\theta_Z)\big)\,\Delta t_i + \psi_Z(t_i,X_i^l,\theta_Z)\,\Delta W_i^l.$$

Loss computing For ease of notation, we define

$$Z_i^l := \psi_Z(t_i, X_i^l, \theta_Z)\big), \qquad f_i^l := f\big(t_i, X_i^l, Y_i^l, \psi_Z^x(t_i, X_i^l, \theta_Z)\big)$$

We also define the local residual r_i^l and the target $\widehat{Z}_i^l$:

$$r_i^l \;=\; Y_{i+1}^l - Y_i^l + f_i^l\,\Delta t_i - Z_i^l\,\Delta W_i^l, \qquad \widehat{Z}_i^l \;=\; \frac{1}{\Delta t_i}\big(Y_{i+1}^l - Y_i^l + f_i^l\,\Delta t_i\big)\,\Big(\Delta W_i^l\Big)^\top$$

The empirical loss is then written :

$$\mathcal{L}(y_0, \theta_Z) \;=\; \mathcal{L}_T(y_0, \theta_Z) \;+\; \lambda_{\text{res}}\,\mathcal{L}_{\text{res}}(y_0, \theta_Z) \;+\; \lambda_Z\,\mathcal{L}_Z(\theta_Z),$$

where $\lambda_{\text{res}}, \lambda_Z > 0$ and :

$$\mathcal{L}_T = \frac{1}{M}\sum_{l=1}^{M}\big\|Y_N^l - g(X_N^l)\big\|^2, \quad \mathcal{L}_{\text{res}} = \frac{1}{M}\sum_{l=1}^{M}\frac{1}{N}\sum_{i=0}^{N-1}\|r_i^l\|^2, \quad \mathcal{L}_Z \;=\; \frac{1}{M}\sum_{l=1}^{M}\frac{1}{N}\sum_{i=0}^{N-1}\|Z_i^l - \widehat{Z}_i^l\|^2$$

where $\big\|.\big\|^2$ denotes the Euclidean norm for vectors and the Frobenius norm for matrices.

Training/Optimization The optimal parameters for (y_0, θ_Z) are given by :

$$(y_0^*, \theta_Z^*) := \underset{(y_0\theta_Z)\in\mathbb{R}\times\mathbb{R}^{d_Z}}{\arg\min}\ \mathcal{L}(y_0, \theta_Z) \tag{2.1}$$

The minimization is done through an Adam iterative optimizer, where λ_{res} and λ_Z can be gradually decreased for better convergence. The parameters' initialization is explained in the following paragraphs.

Generator relaxation We use in this scheme the continuity property of the solutions of a BSDE with regards to the generator as in [49, Proposition 2.1] : we write $g^\delta(t, x, y, z) = \delta g(t, x, y, z)$. We train successively for $\delta_0 = 0 < \delta_1 < \cdots < \delta_K = 1$, reusing (y_0, θ) at each stage, and refine the time grid $t_0 < t_1 < \cdots < t_N$ to reduce time-discretization error. This allows for a simple starting point since $\delta = 0$ corresponds to the heat equation and the continuity of the solutions allows for better convergence towards the desired BSDE.

Warm start For the parameter initialization in the case of $\delta_0 = 0$, we can directly use the theoretical value for y_0 and train for a random initialization for the Z neural network. Numerical results are given in section 2.5.2 of this chapter.

Remark 2.5.6. *$\mathcal{L}_T$ is the original loss function in [26] that only took into consideration terminal time. We add in our scheme a local residual $\mathcal{L}_{\text{res}}$ inspired by [67] to better fit the trajectories to the BSDE. We also add a Malliavin target penalty $\mathcal{L}_Z$ to improve the accuracy of Z, which is used when extracting strategies.*

Remark 2.5.7. *The two additional penalties are not redundant : $\mathcal{L}_{\text{res}}$ enforces the discrete BSDE identity and acts on both Y and Z, whereas $\mathcal{L}_Z$ is a regression and acts only on ψ_Z.*

Remark 2.5.8. *The theoretical Malliavin target is in fact the conditional expectation*

$$Z_i^\star = \frac{1}{\Delta t_i}\,\mathbb{E}\big[(Y_{i+1} - Y_i + f_i \Delta t_i)\,(\Delta W_i)^\top \,\big|\, \mathcal{F}_{t_i}\big].$$

In practice we use the Monte Carlo sample $\widehat{Z}_i^l$ above and minimize an MSE; by the L^2 projection property of condition expectations, the MSE minimizer recovers the conditional expectation without nesting an explicit expectation inside $\widehat{Z}_i^l$. Formally, for a function $h(t_i, X_i)$:

$$\begin{aligned}\mathbb{E}\left[\left\|Z_i^\star - h(t_i, X_i)\right\|^2\right] &= \mathbb{E}\left[\left\|Z_i^\star - \widehat{Z}_i^l + \widehat{Z}_i^l - h(t_i, X_i)\right\|^2\right] \\ &= \mathbb{E}\left[\left\|Z_i^\star - \widehat{Z}_i^l\right\|^2\right] + \mathbb{E}\left[\left\|\widehat{Z}_i^l - h(t_i, X_i)\right\|^2\right]\end{aligned}$$

We can separate the sum since $Z_i^\star = \mathbb{E}\big[\widehat{Z}_i^l, \,\big|\, \mathcal{F}_{t_i}\big]$, which means part of the equation is F_{t_i}-measurable and the other is independant of it.
This means that optimizing a function h using either $Z_i^\star$ or $\widehat{Z}_i^l$ as a target would yield the same minimum.

Numerical results

We give an example in dimension $m = 10$, and using the standard Blotto gain $g(x) = \sum_{i=1}^m \alpha_i \mathbf{1}_{\mathbb{R}^+}(x_i)$ for $\alpha_i = 2 \cdot \mathbf{1}_{i=1} + \mathbf{1}_{2 \le i \le m}$. We use for Z a neural network with two hidden layers of size 400 each, and a SiLU activation function.
For the warm start, we set a constant $\lambda_{\text{res}} = 0$ and $\lambda_Z = 0.06$. Initializing y_0 at the theoretical value of 5.5, and randomly initializing the parameters of Z, we get the following training results for 5000 iteraltions :

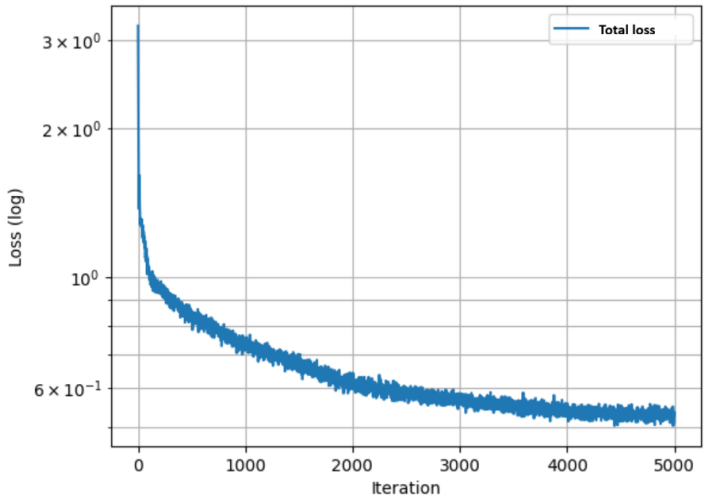


FIGURE 2.1 – Loss by iteration

The final relative error on Y was $1.47.10^{-5}$ and on Z was $6.005.10^{-02}$.

Using the scheme we described, we plot the expected gain of Player 1 as a function of the difference between budgets, using the generator relaxation method.

After training, the optimal strategies can be extracted by selecting the largest component of $Z(0, x_1 - x_2)$, which prescribes the battlefield to prioritize.

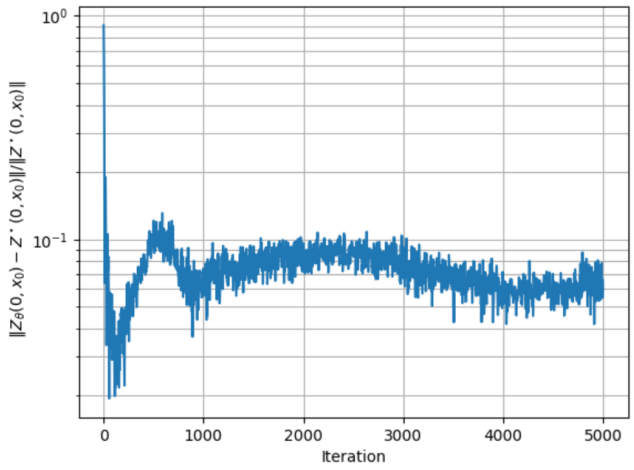


FIGURE 2.2 – Relative error on Z by iteration

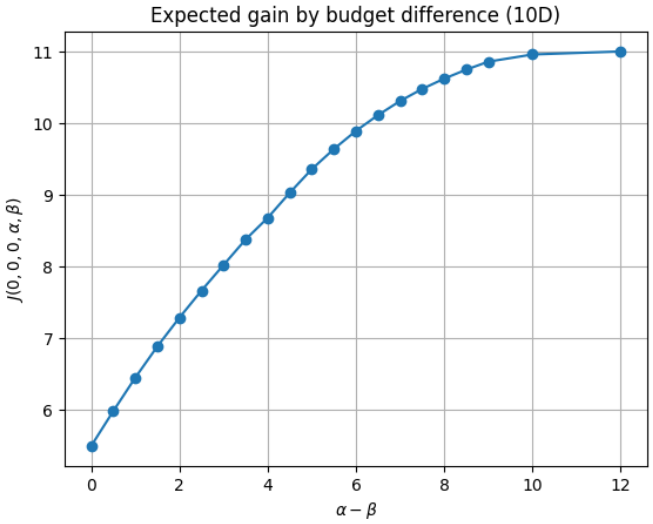


FIGURE 2.3 – Expected gain of Player 1, by budget difference

Chapitre 3

Principal-Agent Blotto game

3.1 Introduction

Principal-agent problems, and contract theory in general, are a important topic in economic theory and game theory, addressing the mathematical modeling that arises when a principal employs one or more agents, whose incentives may not align with their own, to accomplish tasks or jobs. In its classical form, the interaction involves two parties, the principal and the agent, where the principal designs a mechanism to incentivize the agent to act in their best interest through designing a contract.
In this chapter, we introduce a third party (called hacker) that will play against the agent to sabotage the principal. We suppose neither the principal nor the agent know the goal of the hacker, and thus assume a worst-case scenario posture. This will later translate into a second Principal, referred to as the Principal hacker, that will play a zero-sum game against the principal.
From a mathematical viewpoint, we follow the continuous-time dynamic programming and BSDE approach to principal–agent problems developed by [22]. Their framework shows that, under suitable restrictions on admissible contracts, a very general class of models can be reformulated as a standard stochastic control problem for the principal in terms of the agent's continuation utility, using second–order backward SDEs. This methodology already yields explicit solutions in applications to dynamic risk management under moral hazard [21]. Hernández Santibáñez [44] further extends this approach to a principal–multiagent setting in which several competitive agents exert equivalent changes of measure by controlling both the drift and jump characteristics of the output process, and proves the existence of optimal contracts under relatively weak regularity assumptions.

The principal's game. Regarding the principal, they offer the agent a contract at a time horizon T and a continuous budget, both of which are costly. The principal does not observe the action of the agent, so budget embezzlement is possible. The principal then aims to maximize their expected utility by taking into account the optimal strategies that will be played by the agent and the hacker. We now present the game played by the agent and the hacker.

The agent-hacker game. The agent and the hacker play a similar game to the dynamic stochastic Blotto game, introduced in [39]. The game is a time-continuous stochastic resource allocation competition that can be used to model multiple conflicts including cyber-attacks.

Summary of the approach. Within the Stackelberg framework, the principal acts as a leader : by designing the contract offered to the agent, they anticipate the equilibrium response of the agent–hacker game and optimize accordingly. We recall that the Principal and agent do not know the aim of the hacker, so they assume a worst case scenario. In other words, they act as if the hacker is employed by a Principal hacker trying to minimize the principal's utility. The game is rigorously defined in section 3.2, but we can introduce elements here.
Suppose the principal gives the agent a budget process α and a contract ξ_A, while the Principal hacker offers the hacker the contract ξ_H. The agent-hacker game is played on an m-dimensional state process $X_t(a, b)$, where a and b are the controls of the agent and the hacker respectively. Given these parameters, let $J_A(\xi_A, \alpha, a, b)$ and $J_H(\xi_H, a, b)$ be the objective functions of the agent and the hacker respectively. The agent would be maximizing while the hacker minimizing. We first solve the game between the agents by giving a Nash equilibrium $a^\star, b^\star$ which would also be a saddle point, meaning

$$J_A(\xi_A, \alpha, a^\star, b^\star) = \inf_b \sup_a J_A(\xi_A, \alpha, a, b) = \sup_a \inf_b J_A(\xi_A, \alpha, a, b)$$
$$J_H(\xi_A, \alpha, a^\star, b^\star) = \inf_b \sup_a J_H(\xi_H, a, b) = \sup_a \inf_b J_H(\xi_H, a, b)$$

Using the Stackelberg structure, the principal would anticipate this Nash equilibrium to give her contract ξ_A, and assume the worst case scenario, meaning that the principal hacker takes ξ_H to minimize the principal's utility. This worst case scenario problem would be translated mathematically into a principal value function V_P of the form :

$$V_P := \sup_{(\xi_A, \alpha)} \inf_{\xi_H} \mathbb{E}\left[G(X_T(a^\star, b^\star)) - \xi_A - \int_t^T c(\alpha_s)\, \mathrm{d}s \right].$$

where $G : \mathbb{R}^m \to \mathbb{R}$ is a utility function and $c : \mathbb{R} \to \mathbb{R}$ is a cost function.
We then proceed through a representation of the contracts in BSDE form to give a value for the worst case scenario.

Main contributions. This chapter develops a principal–agent version of a stochastic Blotto-type allocation game with an unobserved adversary. First, for a given principal contract ξ_A and budget process α, and for a given hacker contract ξ_H, we characterize pure-strategy Nash equilibria of the continuous-time agent–hacker game via a BSDE representation of continuation utilities ; in the Blotto structure, the equilibrium controls admit explicit feedback forms. Second, since the hacker is hired by a third party and the principal evaluates outcomes under ambiguity about ξ_H, we formalize a worst-case evaluation as an adversarial principal-hacker selecting ξ_H to minimize the principal's value. Under our admissible contract class Ξ, we prove a collapse result showing that the worst-case effect of ξ_H can be eliminated : the principal's robust problem reduces to a zero-sum game against an adverse control b acting on the drift, thereby removing one layer of contract design. Third, in the Markovian case, the collapsed problem is characterized by an Isaacs HJBI equation ; the inner minimization over the adverse control is explicit (a bang–bang minimizer at extreme points) and the principal's remaining pointwise optimization admits a low-dimensional reduction, yielding tractable feedback rules for both the contract and the budget process for standard cost specifications. Finally, we implement a neural-network Deep-BSDE solver for the resulting HJBI/BSDE representation ; exploiting the closed-form inner minimization and the low-dimensional structure of the Hamiltonian, we obtain an efficient numerical procedure and provide simulations of the principal's value as well as the induced budget dynamics.

Structure of the chapter. Section 3.2 formally introduces the setting, and the agent-hacker game in weak formulation, together with the control sets and objectives for the agent, hacker, principal, and principal-hacker. We then solve the agent-hacker game in section 3.3. In section 3.4, we reformulate the problem of the principal by reducing a layer of optimization and then rewrite their value using a PDE through a dynamic programming approach. In section 3.5, we solve the principal's game and give explicit optimal controls. And finally, in section 3.6, we describe the numerical scheme we used and give illustrations.

3.2 Model and Assumptions

The stochastic basis. Throughout the chapter, we fix a finite horizon $T > 0$ and a number $m \in \mathbb{N}^\star$ of battlefields. Let $\Omega = \mathcal{C}([0,T], \mathbb{R}^m)$ be the canonical space of continuous trajectories defined on $[0,T]$ with values in $\mathbb{R}^m$ and let $X : \Omega \to \Omega$ be the canonical process, that is, $X_t(\omega) = \omega(t)$, for every $t \in [0,T]$ and $\omega \in \Omega$. The space Ω is equipped with the topology of uniform convergence, so that it is a complete separable metric space. Let $\mathcal{F}$ be the Borel σ-algebra on Ω, and let P be the Wiener measure on $(\Omega, \mathbb{F})$. We denote by $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ the filtration generated by X and augmented by the P-null sets.

We define, for $\lambda > 0$, the set $A_\lambda := \{x \in \mathbb{R}^m, \forall i = 1, ..., m,\ x_i \ge 0, \sum_{i=1}^m x_i \le \lambda\}$. Such set is where the actions of a player of the Blotto game, with budget λ, take values. We let $B := A_\beta$ where β is a positive constant representing the budget of the hacker, fixed for the rest of the chapter, and supposed to be of public knowledge. Next, let $\mathcal{A}$ be the set of $\mathbb{R}_+$-valued, $\mathbb{F}$-progressively measurable processes and $\mathcal{B}$ be the set of B-valued $\mathbb{F}$-progressively measurable processes. Finally, for any $\alpha \in \mathcal{A}$ representing the budget that the Principal will provide to the Agent, we let $\mathcal{A}(\alpha)$ be the set of $\mathbb{F}$-progressively measurable processes a such that $a_s \in A_{\alpha_s}$ for all $0 \le s \le T$.

We introduce a game in weak formulation, in the sense that the controls of the two players (the agent and the hacker) are exerted via a family of probability measures. Consider any $\alpha \in \mathcal{A}$. For $a \in \mathcal{A}(\alpha)$ and $b \in \mathcal{B}$, we define the probability measure $P^{a,b}$ by a measure change, with the following Radon-Nikodym derivative

$$\frac{dP^{a,b}}{dP} := \exp\left(\int_0^T (a_s - b_s)\mathrm{d}\hat{X}_s - \frac{1}{2}\int_0^T |a_s - b_s|^2 \mathrm{d}s\right).$$

In particular, the process $W_t^{a,b} := X_t - \int_0^t (a_s - b_s)\mathrm{d}s$, $0 \le t \le T$ is a $P^{a,b}$-Brownian motion.

The contract between the principal and the agent. The principal chooses a budget $\alpha \in \mathcal{A}$, which is available for the agent to use in the battlefields. In our model, the principal can not directly observe the actions taken by the agent, so moral hazard is present. As a consequence, the principal creates incentives for the agent in the form of a contract $\xi_A \in \Xi$, where the set of contracts Ξ is defined as the set of $\mathcal{F}_T$-measurable random variables that admit exponential moments of all orders. This means that the principal remunerates the agent at the terminal time T according to the values of the process X.

The agent's and hacker's criteria. Given a contract ξ_A and a budget process α, the agent receives the monetary compensation ξ_A at time T. During the whole period $[0,T]$, he has also the opportunity to divert some of the available budget and make an additional profit, since his

action $a \in \mathcal{A}(\alpha)$ cannot be observed by the principal. We assume that the agent has exponential utility, while the hacker is hired by an external principal and offered a contract ξ_H to sabotage the principal. In the game between the agent and the hacker, the agent's problem is to maximize the criterion J_A while the hacker's is to minimize J_H, both of which are given in definition 3.2.1.

Definition 3.2.1. *Given a contract $\xi_A \in \Xi$ offered to the agent, a budget $\alpha \in \mathcal{A}$ chosen by the principal and a contract $\xi_H \in \Xi$ offered by an external entity to the hacker, we define the game criterion as*

$$J_A(\xi_A, \alpha, a, b) := \mathbb{E}^{P^{a,b}}\left[-\exp\left[-\gamma\left(\xi_A + \int_0^T (\alpha_s - \sum_{i=1}^n (a_s)_i) ds\right)\right]\right],$$
$$J_H(\xi_H, a, b) := \mathbb{E}^{P^{a,b}}\left[\xi_H\right].$$

The best reaction functions of the agent and the hacker are then, respectively :

$$v_A(\xi_A, \alpha, b) := \sup_{a \in \mathcal{A}(\alpha)} J_A(\xi_A, \alpha, a, b), \qquad v_H(\xi_H, a) := \inf_{b \in \mathcal{B}} J_H(\xi_H, a, b).$$

In game theory, a *Nash equilibrium* is the standard concept used to describe the solution of a game. It is a strategy profile, *i.e.*, a control for each player, where neither player benefits from unilaterally deviating.

Definition 3.2.2. *For a given contract $\xi_A \in \Xi$ and budget $\alpha \in \mathcal{A}$ offered to the agent, and a contract $\xi_H \in \Xi$ offered to the hacker, a pair of strategies $(a_0, b_0) \in (\mathcal{A}(\alpha), \mathcal{B})$ is called a Nash equilibrium if :*

- $J_A(\xi_A, \alpha, a, b_0) \leq J_A(\xi_A, \alpha, a_0, b_0)$ *for all* $a \in \mathcal{A}(\alpha)$.
- $J_H(\xi_H, a_0, b_0) \leq J_H(\xi_H, a_0, b)$ *for all* $b \in \mathcal{B}$.

We denote by $NE(\xi_A, \alpha, \xi_H)$ the set of Nash equilibria for this game.

The principal's criterion. The principal offers a budget process $\alpha \in \mathcal{A}$ and a contract $\xi_A \in \Xi$ to the agent, while aiming to maximize her expected payoff by inducing an equilibrium (a, b) between the agent and the hacker. Formally, this would suggest that the principal faces the following problem

$$\sup_{(\xi_A, \alpha) \in \Xi \times \mathcal{A}} \sup_{(a,b) \in NE(\xi_A, \alpha, \xi_H)} \mathbb{E}^{P^{a,b}}\left[G(X_T^{t,x}) - \xi_A - \int_t^T c(\alpha_s)\, \mathrm{d}s\right].$$

where $G : \mathbb{R}^m \to \mathbb{R}$ is a terminal utility function having at most exponential growth (i.e. there exist $r, l > 0$ such that $\forall x \in \mathbb{R}^m$; $G(x) \leq r \exp(l\|x\|)$) and $c : \mathbb{R} \to \mathbb{R}$ is a strictly convex cost function.
However, since the hacker's objective function is unknown, the equilibrium set $NE(\xi_A, \alpha, \xi_H)$ cannot be characterized. To remain robust, we therefore model the hacker as an adversary minimizing the principal's criterion. This is equivalent to having the hacker working as an agent for a external party whose goal is to minimize the principal's objective function. We therefore set the problem with ambiguity on the hacker's contract $\xi_H \in \Xi$. The principal's worst-case scenario problem is then the following

$$V_P := \sup_{(\xi_A, \alpha) \in \Xi \times \mathcal{A}} \inf_{\xi_H \in \Xi} \inf_{(a,b) \in NE(\xi_A, \alpha, \xi_H)} \mathbb{E}^{P^{a,b}}\left[G(X_T) - \xi_A - \int_t^T c(\alpha_s)\, \mathrm{d}s\right].$$

Remark 3.2.3. *We choose to put an* inf *over the Nash equilibria in order to reflect the idea that the principal-hacker has more leverage over which Nash equilibria the agent and the hacker would get to. It is also to be noted that the structure of our problem makes taking the infimum or the supremum irrelevant, because we extract dominant strategies for the game between the agent and the hacker.*

3.3 Solving the game between the agent and the hacker

Let us focus on the Blotto game played by the agent and the hacker, which happens after their contracts have already being offered to them and the budgets have been facilitated. We thus fix $\xi, \xi_H \in \Xi$ and $\alpha \in \mathcal{A}$ for the rest of the section. To solve the game, we use the dynamic programming approach developed in [22]. Start by considering the following 2-dimensional Backward Stochastic Differential Equation (BSDE)

$$Y_t = \xi + \int_t^T F(Z_s, \alpha_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s, \tag{3.1}$$

with the notations $\xi = (\xi_A, \xi_H) \in \Xi^2$, $Z = (Z^1, Z^2) \in \mathbb{R}^{2m}$ and $Z_s \mathrm{d}W_s := (Z_s^1 \cdot \mathrm{d}W_s^1, Z_s^2 \cdot \mathrm{d}W_s^2)^\top$, with $(W_u^i)_u$ a P-Brownian motion. The generator $F : \mathbb{R}^{2m} \times \mathbb{R}_+ \longrightarrow \mathbb{R}^2$ is defined by $F(z, \alpha) := f(z, \alpha, a^\star(z^1, \alpha), b^\star(z^2, \beta))$ with the function $f : \mathbb{R}^{2m} \times \mathbb{R}_+ \times \mathbb{R}^m \times \mathbb{R}^m \longrightarrow \mathbb{R}^2$ defined by

$$f(z, \alpha, a, b) = \begin{bmatrix} f_1(z^1, \alpha, a, b) \\ f_2(z^2, a, b) \end{bmatrix} := \begin{bmatrix} z^1.(a-b) + (\alpha - \sum_{i=1}^m a_i) - \frac{\gamma}{2}||z^1||^2 \\ z^2.(a-b) \end{bmatrix},$$

and the optimizers

$$\begin{cases} a^\star(z^1, \alpha) := \alpha e_i \mathbb{1}_{z_i^1 \geq 1}, \text{ where } i = \arg\max_{i=1,..,m} z^1.e_i, \\ b^\star(z^2, \beta) := \beta e_i \mathbb{1}_{z_i^2 \geq 0}, \text{ where } i = \arg\max_{i=1,..,m} z^2.e_i. \end{cases}$$

In the next result we show that the optimizers above correspond to a saddle point of f, which will represent *a posteriori* (see Proposition 3.3.6) an equilibrium between the agent and the hacker.

Lemma 3.3.1. *For every $\alpha \in \mathbb{R}_+$ and any $(a, b) \in A_\alpha \times B$, we have that*

$$f_1(z^1, \alpha, a^\star(z^1, \alpha), b) = \sup_{a \in A_\alpha} f_1(z^1, \alpha, a, b), \quad f_2(z^2, a, b^\star(z^2, \beta)) = \inf_{b \in B} f_2(z^2, a, b).$$

Proof. Take any $\alpha \in \mathbb{R}_+$, $(a, b) \in A_\alpha \times B$ and $z = (z^1, z^2)^\top \in \mathbb{R}^{2m}$. Let us abuse the notations and write $a^\star = a^\star(z^1, \alpha)$ and $b^\star = b^\star(z^2, \beta)$. We prove for $b^\star$ first. We have from linear optimization

$$f_2(z, a, b) = (a-b).z^2 = a.z^2 - \sum_{i=1}^m b_i (z^2)_i \geq a.z^2 - \beta \Big(\max_{j \in [\![1,m]\!]} (z^2)_j \Big) \mathbb{1}_{\left\{ \max_{j \in [\![1,m]\!]} (z^2)_j \geq 0 \right\}} = f_2(z, \alpha, a, b^\star).$$

We prove now for $a^\star$ in a similar manner. Again, from linear optimization

$$\begin{aligned} f_1(z, \alpha, a, b) &= \alpha - b.z^1 - \frac{\gamma}{2}||z^1||^2 + \sum_{i=1}^m a_i((z^1)_i - 1) \\ &\leq \alpha - b.z^1 - \frac{\gamma}{2}||z^1||^2 + \alpha \Big(\max_{j \in [\![1,m]\!]} (z^1)_j - 1 \Big) \mathbb{1}_{\left\{ \max_{j \in [\![1,m]\!]} (z^1)_j \geq 1 \right\}} = f_1(z, \alpha, a^\star, b). \quad \square \end{aligned}$$

Remark 3.3.2. *Let us interpret the previous result. The processes Z^1 and Z^2 are the sensitivities of the certainty equivalent of the agent and the hacker respectively to their contracts. The hacker's control simply chooses the battlefield with the most sensitivity to put all their budget in. On the other hand, the agent's optimal control is to work on the contract when such work is more profitable than diverting the budget. In that case, the agent also puts all their budget in the most sensitive battlefield.*

Since the work by [22], it is known that in a Principal–Agent interaction the problem of the Principal can be reformulated as a stochastic control problem by restricting, without loss of generality, the set of contracts to terminal values of adequate controlled processes. This result was extended in [27, 44] to the case of multiple agents, provided the weak formulation of the problem involves only equivalent changes of measures. In such a setting, by studying the BSDE associated to the problem, one can characterize the Nash equilibria of the induced game and obtain a convenient parametrization of admissible contracts.

In our setting, we use the BSDE approach in the same spirit as a representation device : we introduce a tractable family of contracts for which the Nash equilibrium between the agent and the hacker can be characterized explicitly. Importantly, we do not apply at this stage the contract-restriction results from the literature to the pair (ξ_A, ξ_H), because in our model the hacker is not hired by the principal but by a third party and the Principal adopts a pessimistic (worst-case) attitude towards the hacker's contract. Applying the reduction directly to (ξ_A, ξ_H) would require a two-sided parametrization in which the components corresponding to ξ_A and ξ_H are controlled by different entities (equivalently, applying the reduction twice), leading to an unnecessarily heavy nested control/game formulation. We avoid this by proving in the next section that the worst-case effect of ξ_H collapses to an adverse choice of the hacker action b. After this collapse, only the Principal's contract ξ_A remains, and we can then apply the reduction in its standard (single-contract) form. Nonetheless, we can follow the main ideas developed in [59, 42] to obtain the following characterization.

Definition 3.3.3. *For any $\alpha \in \mathcal{A}$, we define the admissible set $\mathcal{Z}_{AH}^{\alpha}$ as follows*

$$\mathcal{Z}_{AH}^{\alpha} := \Big\{ Z \text{ predictable } \mathbb{R}^m\text{-valued}: \quad \mathbb{E}\left[\int_0^T \|Z_t\|^2 \mathrm{d}t\right] < \infty,\ \mathbb{E}\left[\int_0^T \big|F(Z_t, \alpha_t)\big| \mathrm{d}t\right] < \infty \Big\}.$$

We also define, for $n \in \mathbb{N}$, the set $\mathcal{Z}_{AH}^{\alpha,n} := \underbrace{\mathcal{Z}_{AH}^{\alpha} \times \cdots \times \mathcal{Z}_{AH}^{\alpha}}_{n \text{ times}}$.

Definition 3.3.4. *For $y = (y^1, y^2) \in \mathbb{R}^2$, $\alpha \in \mathcal{A}$ and $Z \in \mathcal{Z}_{AH}^{\alpha,2}$ we define the following 2-dimensional process*

$$Y_t^{y,Z,\alpha} := y - \int_0^t F(Z_s, \alpha_s)\mathrm{d}s + \int_0^t Z_s \mathrm{d}X_s, \quad t \in [0,T]. \tag{3.2}$$

We denote the coordinates of the process above by $Y^{y,Z,\alpha,1}$ and $Y^{y,Z,\alpha,2}$.

Remark 3.3.5. *The set $\mathcal{Z}_{AH}^{\alpha}$ is chosen to guarantee that the dynamics* (3.2) *are well-defined for all $Z \in \mathcal{Z}_{AH}^{\alpha,2}$, and that $Y^{y,Z,\alpha}$ is an adapted, square-integrable process. Indeed, the imposed conditions ensure both integrals are well-defined and the forward SDE is simply a definition of the process.*

If the contracts of the agent and the hacker have a convenient form (*i.e.* they can be represented as the terminal value of some $Y^{y,Z,\alpha}$), then we can easily characterize the Nash equilibria of the game. The next result can be proved similarly to [44, Corollary 3.31].

Proposition 3.3.6. *Let $y \in \mathbb{R}^2$, $\alpha \in \mathcal{A}$, $Z = (Z^1, Z^2) \in \mathcal{Z}_{AH}^{\alpha,2}$ and the contracts $\xi_A = Y_T^{y,Z,\alpha,1}$, $\xi_H = Y_T^{y,Z,\alpha,2}$. Define the processes $a_s^{\star,Z^1,\alpha} := a^\star(Z_s^1, \alpha_s)$ and $b_s^{\star,Z^2,\beta} := b^\star(Z_s^2, \beta)$, for every $s \in [0,T]$. Then the couple $(a^{\star,Z^1,\alpha}, b^{\star,Z^2,\beta})$ is a Nash equilibrium of the game between the agent and the hacker. Moreover, the value of the agent is given by $v_A(\xi_A, \alpha, b^{\star,Z^2,\beta}) = -e^{-\gamma y^1}$ and the value of the hacker by $v_H(\xi_H, a^{\star,Z^1,\alpha}) = y^2$.*

3.4 Collapsing the Principal Hacker's problem

Given the solution of the game between the agent and the hacker, the principal can now anticipate on it and reformulate their own problem, taking into account the optimal strategies that the players would adopt. The principal and the agent do not know the contract of the hacker, so we remind of the worst-case scenario problem for the principal

$$V_P = \sup_{(\xi_A,\alpha)\in\Xi\times\mathcal{A}} \inf_{\xi_H\in\Xi} \inf_{(a,b)\in NE(\xi_A,\alpha,\xi_H)} \mathbb{E}^{P^{a,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\mathrm{d}s\right].$$

Given how the problem is presented, we can observe that ξ_H is only used to dermine the hacker's control b. So, given a rich enough contract space Ξ, we can actually rewrite the problem and optimize directly over b, as we show next.

Let us first denote :

$$J_P(\xi_A, \alpha; a, b) := \mathbb{E}^{P^{a,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\mathrm{d}s\right]$$

We first need to write Nash equilibria slightly differently for this section, while keeping the same definition.

Definition 3.4.1. *For fixed $(\xi_A, \alpha) \in \Xi \times \mathcal{A}$, define the agent best-response correspondence to a hacker control b :*

$$BR_A(\xi_A, \alpha; b) := \arg\max_{a\in\mathcal{A}(\alpha)} J_A(\xi_A, \alpha; a, b),$$

and for fixed $\xi_H \in \Xi$ the hacker best-response correspondence to an agent control a :

$$BR_H(\xi_H; a) := \arg\min_{b\in B} J_H(\xi_H; a, b).$$

A pair $(a, b) \in \mathcal{A}(\alpha) \times \mathcal{B}$ is a Nash equilibrium for the agent–hacker game if

$$a \in BR_A(\xi_A, \alpha; b) \qquad \text{and} \qquad b \in BR_H(\xi_H; a).$$

The set of Nash equilibria $NE(\xi_A, \alpha, \xi_H)$ is unchanged by this characterization.

Assumption 3.4.2 (Implementability of a worst-case hacker action)**.** *For every (ξ_A, α), there exist $\xi_H^\dagger \in \Xi$ and $(a^\dagger, b^\dagger) \in NE(\xi_A, \alpha, \xi_H^\dagger)$ such that*

$$J_P(\xi_A, \alpha; a^\dagger, b^\dagger) = \inf_{(a,b)\in NE(\xi_A,\alpha,\xi_H^\dagger)} J_P(\xi_A, \alpha; a, b) \quad \text{and} \quad b^\dagger \in \arg\min_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A, \alpha; a, b).$$

Assumption 3.4.2 means that the principal-hacker can implement (via some contract $\xi_H^\dagger$) a hacker action $b^\dagger$ which attains the worst-case value for the principal.

We denote $\mathcal{B}_{\text{ext}}$ the set of $\mathbb{F}$-progressively measurable processes that take values in $B_{\text{ext}} = \{0, \beta e_1, \dots, \beta e_m\}$.

Proposition 3.4.3. *Fix $(\xi_A, \alpha) \in \Xi \times \mathcal{A}$ and assume that the collapsed inner minimization problem admits a minimizer $b^\dagger \in \mathcal{B}_{\text{ext}}$, i.e.*

$$b^\dagger \in \arg\min_{b \in \mathcal{B}} \inf_{a \in BR_A(\xi_A, \alpha; b)} J_P(\xi_A, \alpha; a, b), \qquad B_{\text{ext}} = \{0, \beta e_1, \dots, \beta e_m\}.$$

Define the bounded $\mathbb{F}$-progressively measurable process $Z^{2,\dagger}$ by

$$Z_t^{2,\dagger} := \begin{cases} e_k, & \text{if } b_t^\dagger = \beta e_k \text{ for some } k \in \{1, \dots, m\}, \\ -\mathbf{1}, & \text{if } b_t^\dagger = 0, \end{cases} \qquad 0 \le t \le T.$$

Assume furthermore that ξ_A admits a feedback representation, i.e. there exist $y^1 \in \mathbb{R}$ and $Z^{1,\dagger} \in \mathcal{Z}^\alpha$ such that, with $Z^\dagger := (Z^{1,\dagger}, Z^{2,\dagger})$ and $y := (y^1, y^2)$,

$$\xi_A = Y_T^{y, Z^\dagger, \alpha, 1} \qquad \text{for some } y^2 \in \mathbb{R}.$$

Define the hacker contract

$$\xi_H^\dagger := Y_T^{y, Z^\dagger, \alpha, 2}.$$

Then :

1. *$b^\star(Z_t^{2,\dagger}, \beta) = b_t^\dagger$ for a.e. t, hence $b^\dagger$ is the hacker equilibrium control induced by $\xi_H^\dagger$;*
2. *for any choice of $a^\dagger \in BR_A(\xi_A, \alpha; b^\dagger)$, the pair $(a^\dagger, b^\dagger)$ belongs to $NE(\xi_A, \alpha, \xi_H^\dagger)$;*
3. *$J_P(\xi_A, \alpha; a^\dagger, b^\dagger) = \inf_{(a,b) \in NE(\xi_A, \alpha, \xi_H^\dagger)} J_P(\xi_A, \alpha; a, b)$ for the principal's pessimistic selection.*

Proof. *Proof of point 1.* By definition of $b^\star(\cdot, \beta)$ (Lemma 3.3.1),

$$b^\star(z^2, \beta) = \beta e_i \, \mathbf{1}_{\{\max_j z_j^2 \ge 0\}}, \qquad i \in \arg\max_j z_j^2.$$

If $b_t^\dagger = \beta e_k$, then $Z_t^{2,\dagger} = e_k$ has $\max_j Z_{t,j}^{2,\dagger} = 1 \ge 0$ and unique maximizer k, hence $b^\star(Z_t^{2,\dagger}, \beta) = \beta e_k = b_t^\dagger$. If $b_t^\dagger = 0$, then $Z_t^{2,\dagger} = -\mathbf{1}$ satisfies $\max_j Z_{t,j}^{2,\dagger} = -1 < 0$, hence $b^*(Z_t^{2,\dagger}, \beta) = 0 = b_t^\dagger$. Which means, using $Z_t^{2,\dagger}$ as defined, we can reproduce $b^\dagger$ using the Nash equilibrium strategies we defined in section 3.3.

Proof of point 2. By Proposition 3.3.6 applied to $Z^\dagger = (Z^{1,\dagger}, Z^{2,\dagger})$ and the contracts $(\xi_A, \xi_H^\dagger) = (Y_T^{y,Z^\dagger,\alpha,1}, Y_T^{y,Z^\dagger,\alpha,2})$, the pair

$$\big(a^\star(Z^{1,\dagger}, \alpha),\, b^\star(Z^{2,\dagger}, \beta)\big)$$

is a Nash equilibrium. By point 1, the second component equals $b^\dagger$.

Proof of point 3. Lemma 3.3.1 implies that, for fixed $Z^{2,\dagger}$, the hacker best response $b^*(Z^{2,\dagger}, \beta) = b^\dagger$ minimizes $f_2(z^2, a, b)$ for *any* agent control a. Hence, for the contract $\xi_H^\dagger$,

$$b^\dagger \in BR_H(\xi_H^\dagger; a) \qquad \text{for all admissible } a.$$

Therefore, for any $a^\dagger \in BR_A(\xi_A, \alpha; b^\dagger)$, we have $(a^\dagger, b^\dagger) \in NE(\xi_A, \alpha, \xi_H^\dagger)$. It follows that

$$\inf_{(a,b) \in NE(\xi_A, \alpha, \xi_H^\dagger)} J_P(\xi_A, \alpha; a, b) = \inf_{a \in BR_A(\xi_A, \alpha; b^\dagger)} J_P(\xi_A, \alpha; a, b^\dagger),$$

and choosing $a^\dagger$ as a minimizer in the right-hand side yields the wanted equality. □

This proposition immediately yields the following corollary :

Corollary 3.4.4 (Verification of Assumption 3.4.2)**.** *Fix $(\xi_A, \alpha) \in \Xi \times \mathcal{A}$. Assume that the collapsed inner minimization admits a minimizer $b^\dagger \in \mathcal{B}_{\text{ext}}$ and that the feedback representation condition in Proposition 3.4.3 holds for (ξ_A, α). Then Assumption 3.4.2 holds for (ξ_A, α).*

Collapsed inner problem.

Definition 3.4.5. *For fixed (ξ_A, α), define the principal's inner worst-case value*

$$\mathcal{W}(\xi_A, \alpha) := \inf_{\xi_H \in \Xi} \inf_{(a,b) \in NE(\xi_A, \alpha, \xi_H)} J_P(\xi_A, \alpha; a, b). \tag{3.1}$$

Define the collapsed value where the adverse controller chooses b directly in the full admissible class $\mathcal{B}$:

$$\widetilde{\mathcal{W}}(\xi_A, \alpha) := \inf_{b \in \mathcal{B}} \inf_{a \in BR_A(\xi_A, \alpha; b)} J_P(\xi_A, \alpha; a, b). \tag{3.2}$$

Finally, define the restricted collapsed value

$$\widetilde{\mathcal{W}}_{\text{ext}}(\xi_A, \alpha) := \inf_{b \in \mathcal{B}_{\text{ext}}} \inf_{a \in BR_A(\xi_A, \alpha; b)} J_P(\xi_A, \alpha; a, b), \tag{3.3}$$

where $\mathcal{B}_{\text{ext}}$ is the class of $\mathbb{F}$-progressively measurable B_{ext}-valued processes.

Proposition 3.4.6 (Collapse of the hacker contract)**.** *For every $(\xi_A, \alpha) \in \Xi \times \mathcal{A}$, if Assumption 3.4.2 is verified, then*

$$\mathcal{W}(\xi_A, \alpha) = \widetilde{\mathcal{W}}(\xi_A, \alpha).$$

Consequently, the principal's robust problem

$$V_P := \sup_{(\xi_A, \alpha) \in \Xi \times \mathcal{A}} \mathcal{W}(\xi_A, \alpha)$$

admits the equivalent formulation

$$V_P = \sup_{(\xi_A, \alpha) \in \Xi \times \mathcal{A}} \inf_{b \in \mathcal{B}} \inf_{a \in BR_A(\xi_A, \alpha; b)} \mathbb{E}^{\mathbb{P}^{a,b}} \left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\, ds \right].$$

Proof. Fix (ξ_A, α).

Step 1 : $\mathcal{W}(\xi_A, \alpha) \leq \widetilde{\mathcal{W}}(\xi_A, \alpha)$. By Assumption 3.4.2, there exist $\xi_H^\dagger \in \Xi$ and $(a^\dagger, b^\dagger) \in NE(\xi_A, \alpha, \xi_H^\dagger)$ such that

$$J_P(\xi_A, \alpha; a^\dagger, b^\dagger) = \inf_{(a,b) \in NE(\xi_A, \alpha, \xi_H^\dagger)} J_P(\xi_A, \alpha; a, b) \quad \text{and} \quad b^\dagger \in \arg\min_{b \in \mathcal{B}} \inf_{a \in BR_A(\xi_A, \alpha; b)} J_P(\xi_A, \alpha; a, b).$$

Since $\mathcal{W}(\xi_A, \alpha)$ is an infimum over $\xi_H \in \Xi$, we have

$$\mathcal{W}(\xi_A, \alpha) \leq \inf_{(a,b) \in NE(\xi_A, \alpha, \xi_H^\dagger)} J_P(\xi_A, \alpha; a, b) = J_P(\xi_A, \alpha; a^\dagger, b^\dagger).$$

Moreover, because $b^\dagger$ attains the worst-case value over $\mathcal{B}$, we obtain

$$J_P(\xi_A, \alpha; a^\dagger, b^\dagger) \leq \inf_{b \in \mathcal{B}} \inf_{a \in BR_A(\xi_A, \alpha; b)} J_P(\xi_A, \alpha; a, b) = \widetilde{\mathcal{W}}(\xi_A, \alpha),$$

and therefore $\mathcal{W}(\xi_A, \alpha) \leq \widetilde{\mathcal{W}}(\xi_A, \alpha)$.

Step 2 : $\mathcal{W}(\xi_A, \alpha) \geq \widetilde{\mathcal{W}}(\xi_A, \alpha)$. Let $\xi_H \in \Xi$ be arbitrary and let $(a, b) \in NE(\xi_A, \alpha, \xi_H)$. By definition of Nash equilibrium,

$$a \in BR_A(\xi_A, \alpha; b) \qquad \text{and} \qquad b \in \mathcal{B}.$$

Hence

$$J_P(\xi_A,\alpha;a,b) \geq \inf_{a'\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A,\alpha;a',b) \geq \inf_{b'\in\mathcal{B}} \inf_{a'\in BR_A(\xi_A,\alpha;b')} J_P(\xi_A,\alpha;a',b') = \widetilde{\mathcal{W}}(\xi_A,\alpha).$$

Taking the infimum over $(a,b) \in NE(\xi_A,\alpha,\xi_H)$ and then the infimum over $\xi_H \in \Xi$ yields $\mathcal{W}(\xi_A,\alpha) \geq \widetilde{\mathcal{W}}(\xi_A,\alpha)$.

Combining steps 1 and 2 proves $\mathcal{W}(\xi_A,\alpha) = \widetilde{\mathcal{W}}(\xi_A,\alpha)$. Taking $\sup_{(\xi_A,\alpha)}$ gives the stated reformulation of V_P.

3.5 Solving the Principal's problem

In Section 3.4 we introduced the collapsed inner criterion

$$\widetilde{\mathcal{W}}(\xi_A,\alpha) = \inf_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A,\alpha;a,b),$$

and showed (under Assumption 3.4.2) that, for every (ξ_A,α),

$$\mathcal{W}(\xi_A,\alpha) = \widetilde{\mathcal{W}}(\xi_A,\alpha).$$

In this section, we *start directly from the collapsed principal problem* and derive its HJB characterization. We denote its value by

$$\widetilde{V}_P := \sup_{(\xi_A,\alpha)\in\Xi\times\mathcal{A}} \widetilde{\mathcal{W}}(\xi_A,\alpha) = \sup_{(\xi_A,\alpha)\in\Xi\times\mathcal{A}} \inf_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} \mathbb{E}^{\mathbb{P}^{a,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\,ds\right]. \tag{3.1}$$

At the end of the section, we identify an explicit worst-case control $b^\dagger$ in the Hamiltonian and show that it can be selected B_{ext}-valued. This allows us to apply Proposition 3.4.3 and Corollary 3.4.4 to verify Assumption 3.4.2 in our setting. Combining this verification with Proposition 3.4.6, we conclude that the collapsed value $\widetilde{V}_P$ coincides with the original principal value V_P.

Since the hacker's contract has been eliminated, we apply the contract-restriction machinery in its standard *single-contract* form, i.e. only to the principal–agent contract ξ_A. This yields a stochastic control formulation in terms of the agent's continuation utility and leads to an HJB equation.

3.5.1 Deriving the HJB PDE

We now derive the HJB equation characterizing the collapsed value $\widetilde{V}_P$ defined in (3.1). Throughout, we work under the weak formulation and consider the value of the problem.

Definition 3.5.1. *For $(z,\alpha) \in \mathbb{R}^m \times \mathbb{R}_+$ define*

$$F_0(z,\alpha) := \sup_{a\in\mathcal{A}(\alpha)} \left\{ z\cdot a + \alpha - \mathbf{1}^\top a \right\} - \frac{\gamma}{2}\|z\|^2, \tag{3.2}$$

and define the corresponding explicit best response

$$a^*(z,\alpha) := \alpha\, e_i\, \mathbb{1}_{\{z_i\geq 1\}}, \qquad \text{where } i \in \arg\max_{j=1,\dots,m} z\cdot e_j. \tag{3.3}$$

Equivalently,

$$F_0(z,\alpha) = z\cdot a^*(z,\alpha) + \alpha - \mathbf{1}^\top a^*(z,\alpha) - \frac{\gamma}{2}\|z\|^2. \tag{3.4}$$

Remark 3.5.2. *For fixed (z, α) and any admissible b, the term $-z \cdot b$ is constant in a, hence*

$$\arg\max_{a \in \mathcal{A}(\alpha)} \{z \cdot (a - b) + \alpha - \mathbf{1}^\top a\} = \arg\max_{a \in \mathcal{A}(\alpha)} \{z \cdot a + \alpha - \mathbf{1}^\top a\}.$$

In particular, the feedback control $a^*(z, \alpha)$ defined in (3.3) *belongs to $BR_A(\xi_A, \alpha; b)$ for every b.*

We now introduce the usual admissibility class ensuring that the feedback contract process is well-defined.

Definition 3.5.3. *For any $\alpha \in \mathcal{A}$, define*

$$\mathcal{Z}^\alpha := \Big\{ Z \ \mathbb{F}\text{-progressively measurable } \mathbb{R}^m\text{-valued}: \ \mathbb{E}\Big[\int_0^T \|Z_t\|^2 \, dt\Big] < \infty, \ \mathbb{E}\Big[\int_0^T \big|F_0(Z_t, \alpha_t)\big| \, dt\Big] < \infty \Big\}. \tag{3.5}$$

Remark 3.5.4. *Although b does not appear in the definition of F_0, it enters the performance functionals through $\int_0^T Z_t \cdot b_t \, dt$. This term is automatically integrable uniformly over $b \in \mathcal{B}$ as soon as $Z \in L^2(dt \otimes d\mathbb{P})$. In particular, in the Blotto setting $\|b_t\| \le \beta$ for all admissible b, and for any $Z \in \mathcal{Z}^\alpha$ and any $b \in \mathcal{B}$,*

$$\mathbb{E}\Big[\int_0^T |Z_t \cdot b_t| \, dt\Big] \le \beta\sqrt{T}\Big(\mathbb{E}\int_0^T \|Z_t\|^2 dt\Big)^{1/2} < \infty.$$

Hence we do not need to incorporate b into the definition of $\mathcal{Z}^\alpha$.

Definition 3.5.5 (Feedback contracts after collapse)**.** *For $y \in \mathbb{R}$, $\alpha \in \mathcal{A}$ and $Z \in \mathcal{Z}^\alpha$, define the one-dimensional process*

$$Y_t^{y,Z,\alpha} := y - \int_0^t F_0(Z_s, \alpha_s) \, ds + \int_0^t Z_s \, dX_s, \qquad t \in [0, T], \tag{3.6}$$

and set the principal–agent contract to be $\xi_A := Y_T^{y,Z,\alpha}$.

Remark 3.5.6. *In the weak-formulation framework of [22], one may restrict without loss of generality to principal–agent contracts ξ_A that admit a feedback representation of the form* (3.6)*. In particular, when verifying Assumption 3.4.2 below, we only need to consider such feedback contracts. Moreover, once a bounded worst-case control $b^\dagger \in \mathcal{B}_{\text{ext}}$ has been identified, this feedback restriction provides the required "Z^1-representability" in Proposition 3.4.3 : indeed, for fixed α and bounded $b^\dagger$, one may choose $Z^{1,\dagger}$ so that ξ_A is the first component terminal value in the two-dimensional representation associated with $(Z^{1,\dagger}, Z^{2,\dagger})$ (see the discussion preceding Proposition 3.4.3).*

Let R_0 be the agent's reservation utility. In our setting, the agent's value under the feedback contract $\xi_A = Y_T^{y,Z,\alpha}$ depends on (y, Z, α) only through y (as in Proposition 3.3.6, now with a single component). In particular, the participation constraint is equivalent to $y \ge \bar{y}(R_0)$ and therefore binds at the optimum. With a slight abuse of notation, we denote by R_0 the corresponding reservation level in certainty-equivalent units, so that the binding condition reads $y = R_0$.

Collapsed principal value in feedback form. Using the feedback parametrization above and the best response $a^*(Z,\alpha)$, we can rewrite (3.1) as

$$\widetilde{V}_P := \sup_{\alpha\in\mathcal{A}} \sup_{Z\in\mathcal{Z}^\alpha} \inf_{b\in\mathcal{B}} \mathbb{E}^{\mathbb{P}^{a^*(Z,\alpha),b}}\left[G(X_T) - Y_T^{R_0,Z,\alpha} - \int_0^T c(\alpha_s)\,ds\right]. \tag{3.7}$$

Definition 3.5.7 (Value function)**.** *For $(t,x,y)\in[0,T]\times\mathbb{R}^m\times\mathbb{R}$, we define the collapsed principal value function by*

$$v(t,x,y) := \sup_{\alpha\in\mathcal{A}} \sup_{Z\in\mathcal{Z}^\alpha} \inf_{b\in\mathcal{B}} \mathbb{E}^{\mathbb{P}^{a^*(Z,\alpha),b}}\left[G(X_T) - Y_T^{t,x,y;Z,\alpha} - \int_t^T c(\alpha_s)\,ds\right], \tag{3.8}$$

where, for given controls (α,Z,b), the pair $(X^{t,x},Y^{t,x,y;Z,\alpha})$ is defined as follows :

— *$X^{t,x}$ starts from $X_t^{t,x}=x$ and evolves under $\mathbb{P}^{a^*(Z,\alpha),b}$ according to the weak formulation, i.e. $W^{a,b} := X - \int_0^\cdot (a^*(Z_s,\alpha_s)-b_s)ds$ is a Brownian motion and therefore*

$$dX_s^{t,x} = (a^*(Z_s,\alpha_s)-b_s)\,ds + dW_s^{a,b}, \qquad s\in[t,T];$$

— *$Y^{t,x,y;Z,\alpha}$ starts from $Y_t^{t,x,y;Z,\alpha}=y$ and is given by*

$$Y_s^{t,x,y;Z,\alpha} := y - \int_t^s F_0(Z_r,\alpha_r)\,dr + \int_t^s Z_r\cdot dX_r^{t,x}, \qquad s\in[t,T],$$

so that the terminal contract is $\xi_A = Y_T^{t,x,y;Z,\alpha}$.

Since $Y_T^{y+\delta,Z,\alpha} = Y_T^{y,Z,\alpha}+\delta$ for any $\delta\in\mathbb{R}$, we have the translation property $v(t,x,y+\delta) = v(t,x,y)-\delta$. Hence there exists a function $u:[0,T]\times\mathbb{R}^m\to\mathbb{R}$ such that

$$v(t,x,y) = u(t,x) - y, \qquad \text{and in particular} \qquad \widetilde{V}_P(t,x) = v(t,x,R_0) = u(t,x) - R_0. \tag{3.9}$$

Under the standard assumptions of the single-contract reduction in [22, 44], u is characterized as a viscosity solution of the PDE below.

HJB equation. Formally applying dynamic programming to (3.7) yields an HJB equation for u :

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) + \sup_{\alpha\in A}\sup_{z\in\mathbb{R}^m}\inf_{b\in B}\Big\{(a^*(z,\alpha)-b)\cdot(\nabla u(t,x)-z) + F_0(z,\alpha) - c(\alpha)\Big\} = 0,$$
$$u(T,x) = G(x). \tag{3.10}$$

Using (3.4), we rewrite the term inside braces as

$$a^*(z,\alpha)\cdot\nabla u(t,x) - b\cdot(\nabla u(t,x)-z) + \alpha - \mathbf{1}^\top a^*(z,\alpha) - \frac{\gamma}{2}\|z\|^2 - c(\alpha).$$

Remark 3.5.8. *At this stage, the HJB equation above should be understood at a formal level. Indeed, a fully rigorous derivation based on the dynamic programming principle is delicate in the present setting, since the control spaces are not compact and the standard measurable-selection arguments do not apply directly.*

However, recent work by Hernández-Santibáñez and Mastrolia [43] shows that, for related non-compact control problems, the viscosity characterization can still be obtained through a stochastic

Perron method combined with a comparison argument for the associated PDE. We believe that the same machinery can be adapted to the reduced game considered here. In particular, the explicit form of the reduced Hamiltonian derived below, together with its continuity and its at-most-quadratic growth in the gradient variable, strongly suggests that a stochastic Perron construction and a viscosity comparison result should hold in our setting as well, at least under additional boundedness assumptions on the terminal condition.

Accordingly, in the present manuscript, we use the HJB equation as a formal guide and focus on the structural consequences that follow from its Hamiltonian analysis.

Identification of the worst-case control and verification of the collapse.

Definition 3.5.9. *Fix $(\xi_A, \alpha) \in \Xi \times \mathcal{A}$ and assume that α is essentially bounded, i.e. $\|\alpha\|_\infty < \infty$. Assume that ξ_A is given in feedback form by Definition 3.5.5 :*

$$\xi_A \;=\; Y_T^{R_0, Z^A, \alpha},$$

for some $Z^A \in \mathcal{Z}^\alpha$, and define the associated agent best response

$$a_t^* := a^*(Z_t^A, \alpha_t), \qquad 0 \le t \le T.$$

Define the adapted process

$$\kappa_t \;:=\; F_0(Z_t^A, \alpha_t) - c(\alpha_t) \;=\; Z_t^A \cdot a_t^* + \alpha_t - \mathbf{1}^\top a_t^* - \frac{\gamma}{2}\|Z_t^A\|^2 - c(\alpha_t), \qquad 0 \le t \le T.$$

For any admissible hacker control $b \in \mathcal{B}$, define the generator

$$f^b(t,z) \;:=\; (a_t^* - b_t) \cdot (z - Z_t^A) \;+\; \kappa_t, \qquad (t,z) \in [0,T] \times \mathbb{R}^m.$$

Using (3.4)*, this can be rewritten as*

$$f^b(t,z) = a_t^* \cdot z - b_t \cdot (z - Z_t^A) + \alpha_t - \mathbf{1}^\top a_t^* - \frac{\gamma}{2}\|Z_t^A\|^2 - c(\alpha_t).$$

Consider the corresponding BSDE under the reference Wiener measure $\mathbb{P}$ (so that X is the canonical Brownian motion)

$$Y_t^b = G(X_T) + \int_t^T f^b(s, Z_s^b)\, ds - \int_t^T Z_s^b \cdot dX_s, \qquad 0 \le t \le T. \tag{3.11}$$

Define also the envelope generator

$$f(t,z) \;:=\; \inf_{b \in B} \left\{ (a_t^* - b) \cdot (z - Z_t^A) \right\} + \kappa_t$$

that is,

$$f(t,z) = a_t^* \cdot (z - Z_t^A) \;-\; \beta \max\left(0, \max_{1 \le i \le m} (z_i - (Z_t^A)_i)\right) \;+\; \kappa_t,$$

or equivalently,

$$f(t,z) = a_t^* \cdot z - \beta \max\left(0, \max_{1 \le i \le m} (z_i - (Z_t^A)_i)\right) + \alpha_t - \mathbf{1}^\top a_t^* - \frac{\gamma}{2}\|Z_t^A\|^2 - c(\alpha_t).$$

Define the envelope BSDE

$$Y_t = G(X_T) + \int_t^T f(s, Z_s)\, ds - \int_t^T Z_s \cdot dX_s, \qquad 0 \le t \le T. \tag{3.12}$$

Definition 3.5.10. *In the setting of Definition 3.5.9, let (Y, Z) be a solution of (3.12). Define the progressively measurable B_{ext}-valued process $b^\dagger$ by*

$$b_t^\dagger := \begin{cases} 0, & \text{if } \max_{1\le i\le m}\left(Z_{t,i} - (Z_t^A)_i\right) \le 0, \\ \beta e_{i^*(t)}, & \text{otherwise, where } i^*(t) \in \arg\max_{1\le i\le m}\left(Z_{t,i} - (Z_t^A)_i\right) \text{ (tie-break by smallest index).} \end{cases}$$

Proposition 3.5.11 (Construction of a B_{ext}-valued worst-case control)**.** *Under the assumptions of Definition 3.5.9 :*

1. *For every $b \in \mathcal{B}$, the BSDE (3.11) admits a unique solution (Y^b, Z^b), and the envelope BSDE (3.12) admits a unique solution (Y, Z).*
2. *For every $b \in \mathcal{B}$, the initial value satisfies*

$$Y_0^b - R_0 \;=\; \mathbb{E}^{\mathbb{P}^{a^*,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\,ds\right] \;=\; J_P(\xi_A, \alpha; a^*, b).$$

3. *The process $b^\dagger$ defined in Definition 3.5.10 belongs to $\mathcal{B}_{\text{ext}}$ and is a global minimizer of the collapsed inner problem for the fixed pair (ξ_A, α) :*

$$J_P(\xi_A, \alpha; a^*, b^\dagger) \;=\; \inf_{b\in\mathcal{B}} J_P(\xi_A, \alpha; a^*, b) \;=\; \inf_{b\in\mathcal{B}} \inf_{a\in BR_A(\xi_A,\alpha;b)} J_P(\xi_A, \alpha; a, b).$$

Proof. *Proof of point 1.* Fix $(\xi_A, \alpha) \in \Xi \times \mathcal{A}$. Since $\|a_t^*\| \le \alpha_t \le \|\alpha\|_\infty$ and $\|b_t\| \le \beta$, the map $z \mapsto f^b(t, z)$ is globally Lipschitz with Lipschitz constant bounded by $\|\alpha\|_\infty + \beta$, uniformly in (t, ω). Moreover $f(t, \cdot)$ is also globally Lipschitz with the same bound. Since G has at most exponential growth, $G(X_T)$ is L^2, and classical Lipschitz BSDE theory yields existence and uniqueness of solutions to (3.11) and (3.12), as well as the comparison principle ; see [63].

Proof of point 2. Fix $b \in \mathcal{B}$. Under $\mathbb{P}^{a^*,b}$ we have the weak dynamics

$$dX_t = (a_t^* - b_t)\,dt + dW_t^{a^*,b}, \qquad 0 \le t \le T.$$

Recall that ξ_A is in feedback form :

$$\xi_A = Y_T^{R_0, Z^A, \alpha} = R_0 - \int_0^T F_0(Z_s^A, \alpha_s)\,ds + \int_0^T Z_s^A \cdot dX_s,$$

and

$$\kappa_t := F_0(Z_t^A, \alpha_t) - c(\alpha_t), \qquad 0 \le t \le T.$$

Starting from the BSDE (3.11) under $\mathbb{P}$,

$$Y_t^b = G(X_T) + \int_t^T \Big((a_s^* - b_s)\cdot(Z_s^b - Z_s^A) + \kappa_s\Big)\,ds - \int_t^T Z_s^b \cdot dX_s,$$

we rewrite the stochastic integral under $\mathbb{P}^{a^*,b}$ using $dX_s = (a_s^* - b_s)\,ds + dW_s^{a^*,b}$:

$$\int_t^T Z_s^b \cdot dX_s = \int_t^T Z_s^b \cdot (a_s^* - b_s)\,ds + \int_t^T Z_s^b \cdot dW_s^{a^*,b}.$$

Substituting and simplifying gives, $\mathbb{P}^{a^*,b}$-a.s.,

$$Y_t^b = G(X_T) + \int_t^T \Big(\kappa_s - Z_s^A \cdot (a_s^* - b_s)\Big)\,ds - \int_t^T Z_s^b \cdot dW_s^{a^*,b}.$$

Using (3.4), the drift term can be rewritten as

$$\kappa_s - Z_s^A \cdot (a_s^* - b_s) = \alpha_s - \mathbf{1}^\top a_s^* - \frac{\gamma}{2}\|Z_s^A\|^2 - c(\alpha_s) + Z_s^A \cdot b_s.$$

Therefore,

$$Y_t^b = \mathbb{E}^{\mathbb{P}^{a^*,b}}\left[G(X_T) + \int_t^T \Big(\kappa_s - Z_s^A \cdot (a_s^* - b_s)\Big)\, ds \;\Big|\; \mathcal{F}_t\right].$$

On the other hand, using the feedback expression for ξ_A and again $dX_s = (a_s^* - b_s)\, ds + dW_s^{a^*,b}$, we have

$$\int_0^T Z_s^A \cdot dX_s = \int_0^T Z_s^A \cdot (a_s^* - b_s)\, ds + \int_0^T Z_s^A \cdot dW_s^{a^*,b}.$$

Hence,

$$\begin{aligned} G(X_T) - \xi_A - \int_0^T c(\alpha_s)\, ds &= G(X_T) - R_0 + \int_0^T \big(F_0(Z_s^A, \alpha_s) - c(\alpha_s)\big)\, ds - \int_0^T Z_s^A \cdot dX_s \\ &= G(X_T) - R_0 + \int_0^T \Big(\kappa_s - Z_s^A \cdot (a_s^* - b_s)\Big)\, ds - \int_0^T Z_s^A \cdot dW_s^{a^*,b}. \end{aligned}$$

Taking $\mathbb{P}^{a^*,b}$-conditional expectations, the stochastic integral has mean zero, and we obtain for all $t \in [0,T]$,

$$\mathbb{E}^{\mathbb{P}^{a^*,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\, ds \;\Big|\; \mathcal{F}_t\right] = Y_t^b - R_0.$$

In particular, at $t = 0$ this yields

$$Y_0^b - R_0 = \mathbb{E}^{\mathbb{P}^{a^*,b}}\left[G(X_T) - \xi_A - \int_0^T c(\alpha_s)\, ds\right] = J_P(\xi_A, \alpha; a^*, b).$$

Proof of point 3. First, by Remark 3.5.2, $a^* = a^*(Z^A, \alpha)$ belongs to $BR_A(\xi_A, \alpha; b)$ for every b, hence

$$\inf_{a \in BR_A(\xi_A,\alpha;b)} J_P(\xi_A, \alpha; a, b) = J_P(\xi_A, \alpha; a^*, b) \qquad \text{for all } b \in \mathcal{B}.$$

Next, by construction of the envelope generator, for every (t, z) and every admissible b,

$$f(t,z) \le f^b(t,z),$$

and by definition of $b^\dagger$ we have $f(t, Z_t) = f^{b^\dagger}(t, Z_t)$ for a.e. t. Therefore the unique solution (Y, Z) of (3.12) also solves (3.11) with $b = b^\dagger$, so $Y_0 = Y_0^{b^\dagger}$. By the BSDE comparison principle [63],

$$Y_0 \le Y_0^b \qquad \text{for all } b \in \mathcal{B},$$

hence $Y_0^{b^\dagger} \le Y_0^b$ for all b. Using point 2, this yields

$$J_P(\xi_A, \alpha; a^*, b^\dagger) = \inf_{b \in \mathcal{B}} J_P(\xi_A, \alpha; a^*, b) = \inf_{b \in \mathcal{B}} \inf_{a \in BR_A(\xi_A,\alpha;b)} J_P(\xi_A, \alpha; a, b),$$

and $b^\dagger \in \mathcal{B}_{\text{ext}}$ by construction. □

Proposition 3.5.12. *In the Principal–Agent Blotto game, the collapsed and original values coincide :*

$$\widetilde{V}_P = V_P.$$

Proof. In (3.10), the adverse control b enters only through the linear term

$$-b \cdot (\nabla u(t,x) - z).$$

Since the pointwise constraint set B is the Blotto simplex, the inner minimization is a linear program and is therefore attained at an extreme point. Consequently,

$$\inf_{b\in B}\{-b\cdot q\} = \inf_{b\in B_{\text{ext}}}\{-b\cdot q\} = -\beta \max\Big(0, \max_{1\le i\le m} q_i\Big), \qquad q := \nabla u(t,x) - z.$$

While this identifies the pointwise worst-case control in the Isaacs Hamiltonian, we still need a worst-case adverse *process* for the collapsed inner problem. For fixed (ξ_A, α) (with α bounded), this is provided by Proposition 3.5.11, which constructs a B_{ext}-valued progressively measurable control $b^\dagger \in \mathcal{B}_{\text{ext}}$ attaining the infimum in the collapsed inner minimization.

By Remark 3.5.6, we may restrict without loss of generality to principal–agent contracts ξ_A that admit a feedback representation. Fix such a pair (ξ_A, α) and let $b^\dagger \in \mathcal{B}_{\text{ext}}$ be a worst-case minimizer in the collapsed inner problem. Define $Z^{2,\dagger}$ by

$$Z_t^{2,\dagger} := \begin{cases} e_k, & \text{if } b_t^\dagger = \beta e_k \text{ for some } k \in \{1,\dots,m\}, \\ -\mathbf{1}, & \text{if } b_t^\dagger = 0, \end{cases} \qquad 0 \le t \le T,$$

so that $b^\star(Z_t^{2,\dagger}, \beta) = b_t^\dagger$ for a.e. t (Lemma 3.3.1).

To apply Proposition 3.3.6 with the choice of $Z^{2,\dagger}$ encoding $b^\dagger$, we need a process $Z^{1,\dagger}$ such that ξ_A can be written as the first-component terminal value in the corresponding two-dimensional feedback representation. This is obtained by solving the one-dimensional quadratic BSDE

$$Y_t^{1,\dagger} = \xi_A + \int_t^T \Big(F_0(Z_s^{1,\dagger}, \alpha_s) - Z_s^{1,\dagger}\cdot b_s^\dagger\Big)\, ds - \int_t^T Z_s^{1,\dagger}\cdot dX_s, \qquad 0 \le t \le T. \tag{3.13}$$

Since $b^\dagger$ is bounded and the generator has at most quadratic growth and is continuous in $Z^{1,\dagger}$, standard results on quadratic BSDEs ensure that (3.13) admits a solution under our integrability class Ξ; see [14]. Setting $y^1 := Y_0^{1,\dagger}$ and $Z^\dagger := (Z^{1,\dagger}, Z^{2,\dagger})$ then yields the required representation $\xi_A = Y_T^{y,Z^\dagger,\alpha,1}$ in Proposition 3.3.6.

Defining $\xi_H^\dagger := Y_T^{y,Z^\dagger,\alpha,2}$ then yields, by Proposition 3.4.3 and Corollary 3.4.4, a Nash equilibrium $(a^\dagger, b^\dagger) \in NE(\xi_A, \alpha, \xi_H^\dagger)$ which realizes the principal's pessimistic selection among equilibria and verifies Assumption 3.4.2.

Combining this verification with Proposition 3.4.6 yields $\widetilde{V}_P = V_P$, i.e. the collapsed and original values coincide. □

3.5.2 Pointwise optimization in the Hamiltonian

Definition 3.5.13. *Fix $(t,x) \in [0,T]\times\mathbb{R}^m$ and set*

$$p := \nabla u(t,x) \in \mathbb{R}^m, \qquad M(p-z) := \max\Big(0, \max_{1\le i\le m}(p_i - z_i)\Big), \qquad B_{\text{ext}} := \{0, \beta e_1, \dots, \beta e_m\}.$$

For $z \in \mathbb{R}^m$, define the index selector

$$i(z) := \min \arg\max_{1\le j\le m} z_j,$$

and recall the agent best response

$$a^*(z,\alpha) := \alpha\, e_{i(z)}\, \mathbf{1}_{\{z_{i(z)}\geq 1\}}.$$

Also set

$$F_0(z,\alpha) := z\cdot a^*(z,\alpha) + \alpha - \mathbf{1}^\top a^*(z,\alpha) - \frac{\gamma}{2}\|z\|^2.$$

Definition 3.5.14. *For $q\in\mathbb{R}^m$, define the bang–bang control*

$$b^*(q) := \begin{cases} 0, & \max_{1\leq i\leq m} q_i \leq 0, \\ \beta e_{i^*(q)}, & \text{otherwise, where } i^*(q)\in \arg\max_{1\leq i\leq m} q_i \text{ (tie-break by smallest index).} \end{cases}$$

Definition 3.5.15 (One-dimensional epigraph reduction in t)**.** *For $p,z\in\mathbb{R}^m$, set*

$$M(p-z) = \min\Big\{t\geq 0:\ p_i - z_i \leq t\ \ \forall i\Big\}.$$

For $t\geq 0$, define the minimal-norm feasible choice

$$z_i(t) := (p_i - t)_+, \qquad i=1,\dots,m.$$

For each $k\in\{1,\dots,m\}$, define

$$\mathcal{R}_k := \{z\in\mathbb{R}^m:\ z_k = \max_j z_j,\ z_k\geq 1\}, \qquad z_i^{(k)}(t) := (p_i - t)_+,\ i\neq k, \qquad z_k^{(k)}(t) := \max\{1, p_{\max} - t\},$$

where $p_{\max} := \max_i p_i$.

Definition 3.5.16 (One-slope maximization in α)**.** *Let $c:\mathbb{R}_+\to\mathbb{R}$ be the instantaneous cost. For $\theta\in\mathbb{R}$, define*

$$\Psi(\theta) := \sup_{\alpha\geq 0}\{\alpha\theta - c(\alpha)\},$$

and let $\alpha^*(\theta)$ denote any maximizer (when it exists).

Proposition 3.5.17 (Pointwise optimal controls in the Isaacs Hamiltonian)**.** *Work with the notation of Definitions 3.5.13–3.5.16. Consider the Hamiltonian term in* (3.14) *at gradient $p = \nabla u(t,x)$.*

1. ***Worst-case hacker action.*** *For any (α, z),*

$$\inf_{b\in B}\{-b\cdot(p-z)\} = \inf_{b\in B_{\text{ext}}}\{-b\cdot(p-z)\} = -\beta\, M(p-z),$$

   and the minimizer is given by $b^*(p-z)$ in Definition 3.5.14.

2. ***Agent best response.*** *For any (α,z), the agent best response is $a^*(z,\alpha)$ given in Definition 3.5.13, and the corresponding driver is $F_0(z,\alpha)$.*

3. ***Optimization over z for fixed α.*** *After eliminating b, the pointwise Hamiltonian reduces to* (3.16)*. For each fixed $\alpha\geq 0$, the maximization over z splits into :*
   - *a* no-investment *branch $\max_i z_i < 1$, whose value is given by* (3.21) *and whose optimizer is $z^{(0)} = z(t_0)$ with t_0 as defined below* (3.21) *;*
   - investment *branches on $\mathcal{R}_k$ ($k=1,\dots,m$), whose values are given by* (3.24) *and whose optimizers are $z^{(k)}(t_k)$, where t_k attains the infimum in* (3.24)*.*

4. ***Reduction to one-dimensional maximization in α.*** *After optimizing over z, the Hamiltonian takes the form*

$$H(p) = \sup_{\alpha \geq 0} \max\Big\{ H_0(p;\alpha),\ \max_{1\leq k\leq m} H_k(p;\alpha) \Big\},$$

with H_0, H_k given by (3.26)–(3.27)*. Equivalently,*

$$H(p) = \max\Big\{ K_0(p) + \Psi(1),\ \max_{1\leq k\leq m} \big(K_k(p) + \Psi(p_k)\big) \Big\},$$

where

$$K_0(p) := -\inf_{t \geq \max(0, p_{\max}-1)} \left\{ \frac{\gamma}{2} \sum_{i=1}^{m} (p_i - t)_+^2 + \beta t \right\},$$

and

$$K_k(p) := -\inf_{t\geq 0} \left\{ \frac{\gamma}{2} \Big((z_k^{(k)}(t))^2 + \sum_{i\neq k} (p_i - t)_+^2 \Big) + \beta t \right\}.$$

5. ***Closed-form maximization in α for two standard costs.*** *Let Ψ and $\alpha^*(\cdot)$ be as in Definition 3.5.16. In the branch H_0 the slope is $\theta = 1$, while in the branch H_k the slope is $\theta_k = p_k$.*
   - *If $c(\alpha) = \frac{\lambda}{2}\alpha^2$, then $\alpha^*(\theta) = \theta_+/\lambda$ and $\Psi(\theta) = \theta_+^2/(2\lambda)$.*
   - *If $c(\alpha) = \lambda_1 e^{\lambda_2 \alpha}$, then $\alpha^*(\theta)$ and $\Psi(\theta)$ are obtained by maximizing $\alpha\theta - \lambda_1 e^{\lambda_2\alpha}$ over $\alpha \geq 0$.*

With the convention $Y = y - \int_0^{\cdot} F_0\, ds + \int_0^{\cdot} Z\, dX$, the HJB equation reads

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) + \sup_{\alpha\geq 0} \sup_{z\in\mathbb{R}^m} \inf_{b\in B} \Big\{ (a^*(z,\alpha) - b)\cdot p + F_0(z,\alpha) - z\cdot\big(a^*(z,\alpha) - b\big) - c(\alpha) \Big\} = 0, \qquad u(T,x) = G(x), \tag{3.14}$$

and using (3.4) one checks that the term inside braces can be rewritten as

$$a^*(z,\alpha)\cdot p \ - \ b\cdot(p-z) \ + \ \alpha - \mathbf{1}^\top a^*(z,\alpha) \ - \ \frac{\gamma}{2}\|z\|^2 \ - \ c(\alpha). \tag{3.15}$$

Set $I(z) := \mathbf{1}_{\{\max_i z_i \geq 1\}} = \mathbf{1}_{\{z_{i(z)}\geq 1\}}$. Then we have $a^*(z,\alpha) = 0$ on $\{I(z) = 0\}$ and $a^*(z,\alpha) = \alpha e_{i(z)}$ on $\{I(z) = 1\}$. The Hamiltonian then becomes

$$\sup_{\alpha\geq 0} \sup_{z\in\mathbb{R}^m} \Big\{ \alpha(1 - I(z)) \ + \ \alpha I(z)\, p_{i(z)} \ - \ c(\alpha) \ - \ \frac{\gamma}{2}\|z\|^2 \ - \ \beta\, M(p-z) \Big\}. \tag{3.16}$$

The rest of this section details the optimization.

Optimization over z for a fixed α

We first note that the max-penalty $M(p-z)$ admits the representation

$$M(p-z) = \min\Big\{ t \geq 0 :\ p_i - z_i \leq t \ \ \forall i \Big\}. \tag{3.17}$$

Consequently, for any function $\Phi(z)$,

$$\sup_{z\in\mathbb{R}^m} \Big\{ \Phi(z) - \beta M(p-z) \Big\} = \sup_{t\geq 0} \sup_{\substack{z\in\mathbb{R}^m \\ p_i - z_i \leq t,\ \forall i}} \Big\{ \Phi(z) - \beta t \Big\}. \tag{3.18}$$

In particular, the constraint $p_i - z_i \le t$ is equivalent to $z_i \ge p_i - t$ for all i. Since the only strictly concave term in our objective is the quadratic penalty $-\frac{\gamma}{2}\|z\|^2$, the inner maximization over z for a fixed t is achieved by the *minimal-norm* feasible choice,

$$z_i(t) := (p_i - t)_+, \qquad i = 1, \dots, m. \tag{3.19}$$

Case 1 : $\max_i z_i < 1$. On $\{\max_i z_i < 1\}$ we have $I(z) = 0$, and for fixed α the z-problem reduces to

$$\sup_{\max_i z_i < 1} \Big\{ -\frac{\gamma}{2}\|z\|^2 - \beta\, M(p - z) \Big\}. \tag{3.20}$$

Using (3.18) with $\Phi(z) = -\frac{\gamma}{2}\|z\|^2$ and the projection formula (3.19), we obtain

$$\sup_{\max_i z_i < 1} \Big\{ -\frac{\gamma}{2}\|z\|^2 - \beta\, M(p - z) \Big\} = -\inf_{t \ge 0} \Big\{ \frac{\gamma}{2} \sum_{i=1}^{m} (p_i - t)_+^2 + \beta t \Big\},$$

where the constraint $\max_i z_i < 1$ becomes $\max_i (p_i - t)_+ < 1$. Writing $p_{\max} := \max_i p_i$, this is equivalent to $t \ge \max(0, p_{\max} - 1)$. Therefore,

$$\sup_{\max_i z_i < 1} \Big\{ -\frac{\gamma}{2}\|z\|^2 - \beta\, M(p - z) \Big\} = -\inf_{t \ge \max(0, p_{\max} - 1)} \left\{ \frac{\gamma}{2} \sum_{i=1}^{m} (p_i - t)_+^2 + \beta t \right\}. \tag{3.21}$$

We can now give an explicit optimizer as a function of p : let $p_{(1)} \ge \cdots \ge p_{(m)}$ be the nonincreasing rearrangement of $(p_1, \dots, p_m)$ and set $L := \max(0, p_{(1)} - 1)$. For $k = 1, \dots, m$ define

$$t_k := \frac{1}{k} \left(\sum_{i=1}^{k} p_{(i)} - \frac{\beta}{\gamma} \right).$$

There exists a unique index $k^\star \in \{1, \dots, m\}$ such that $t_{k^\star} \in [p_{(k^\star+1)}, p_{(k^\star)}]$ (with the convention $p_{(m+1)} := -\infty$), and the unconstrained minimizer of $t \mapsto \frac{\gamma}{2} \sum_{i=1}^{m} (p_i - t)_+^2 + \beta t$ is $t_{k^\star}$. Consequently, the constrained minimizer is

$$t_0 = \max\{L,\ t_{k^\star}\}.$$

Case 2 : $\max_i z_i \ge 1$. Fix $k \in \{1, \dots, m\}$ and consider the region

$$\mathcal{R}_k := \{ z \in \mathbb{R}^m :\ z_k = \max_j z_j,\ z_k \ge 1 \},$$

so that on $\mathcal{R}_k$ we have $i(z) = k$ and $I(z) = 1$. For fixed α, the corresponding z-problem is

$$\sup_{z \in \mathcal{R}_k} \Big\{ -\frac{\gamma}{2}\|z\|^2 - \beta\, M(p - z) \Big\}. \tag{3.22}$$

Using the reformulation (3.18) with

$$\Phi_k(z) := -\frac{\gamma}{2}\|z\|^2,$$

we rewrite (3.22) as

$$\sup_{t \ge 0} \sup_{\substack{z \in \mathcal{R}_k \\ p_i - z_i \le t,\ \forall i}} \Big\{ -\frac{\gamma}{2}\|z\|^2 - \beta t \Big\}.$$

For fixed t, the constraints $p_i - z_i \le t$ amount to $z_i \ge p_i - t$ for all i ; minimizing the norm under these constraints and the additional conditions defining $\mathcal{R}_k$ yields :

$$z_i^{(k)}(t) := (p_i - t)_+, \qquad i \neq k, \qquad z_k^{(k)}(t) := \max\Big\{1,\ \max_{1\le j\le m}(p_j - t)_+\Big\} = \max\{1, p_{\max} - t\}. \tag{3.23}$$

Substituting $z^{(k)}(t)$ into the objective gives the one-dimensional reduction

$$\sup_{z\in\mathcal{R}_k}\Big\{-\frac{\gamma}{2}\|z\|^2 - \beta\, M(p-z)\Big\} = -\inf_{t\ge 0}\Big\{\frac{\gamma}{2}\Big((z_k^{(k)}(t))^2 + \sum_{i\neq k}(p_i - t)_+^2\Big) + \beta t\Big\}. \tag{3.24}$$

An optimizer on $\mathcal{R}_k$ is given by $z^{(k)}(t_k)$ where t_k attains the infimum.

Combining (3.21) and (3.24), the $\sup_{z\in\mathbb{R}^m}$ in (3.16) can be written, for each fixed $\alpha \ge 0$, as the maximum of :

— the no-investment branch $\max_i z_i < 1$, with optimizer $z^{(0)} = z(t_0)$ given by (3.19) ;

— the investment branches $\mathcal{R}_k$, with optimizers $z^{(k)}(t_k)$ given by (3.23), $k = 1, \dots, m$.

Reduction to a one-dimensional maximization in α

After optimizing over z as above, the Hamiltonian at gradient $p = \nabla u(t,x)$ takes the form

$$H(p) = \sup_{\alpha\ge 0}\ \max\Big\{H_0(p;\alpha),\ \max_{1\le k\le m} H_k(p;\alpha)\Big\}, \tag{3.25}$$

where

$$H_0(p;\alpha) := \alpha - c(\alpha)\ -\ \inf_{t\ge \max(0,p_{\max}-1)}\left\{\frac{\gamma}{2}\sum_{i=1}^m (p_i - t)_+^2 + \beta t\right\}, \tag{3.26}$$

$$H_k(p;\alpha) := \alpha p_k - c(\alpha)\ -\ \inf_{t\ge 0}\left\{\frac{\gamma}{2}\Big((z_k^{(k)}(t))^2 + \sum_{i\neq k}(p_i - t)_+^2\Big) + \beta t\right\}, \tag{3.27}$$

with $z_k^{(k)}(t)$ given by (3.23).

The representation (3.25) highlights a decomposition into two regimes. The term $H_0(p;\alpha)$ corresponds to a *no-investment* regime : incentives remain below the activation threshold, so the agent is not induced to concentrate the budget on a single battlefield, and the adversarial effect reduces to the penalty $\beta\, t$. The scalar t admits a clear interpretation as a *capping level* for the quantity $p - z$: minimizing over t amounts to choosing the smallest sabotage compatible with a quadratic cost $\|z\|^2$, which yields the *water-filling* structure $z_i = (p_i - t)_+$.

In contrast, the terms $H_k(p;\alpha)$ describe an *investment* regime : incentives are strong enough to trigger a concentrated allocation on battlefield k (through the constraint $z_k \ge 1$), and the principal compares the marginal gain αp_k to the instantaneous cost $c(\alpha)$, while still paying both the term βt and the quadratic penalty in z. After optimizing over z, the remaining pointwise optimization is one-dimensional in α.

Thus, after the z-optimization, the remaining pointwise optimization is one-dimensional in α, and its explicit form depends on the choice of the instantaneous cost c. In the next sections, we give two standard examples.

Quadratic instantaneous cost

We first solve for a quadratic cost function

$$c(\alpha) = \frac{\lambda}{2}\alpha^2, \qquad \lambda > 0.$$

For $\theta \in \mathbb{R}$, define the one-slope maximization problem

$$\Psi_{\text{quad}}(\theta) := \sup_{\alpha \geq 0} \Big\{ \alpha\, \theta - \frac{\lambda}{2}\alpha^2 \Big\}. \tag{3.28}$$

A direct computation yields

$$\alpha_{\text{quad}}(\theta) = \frac{\theta_+}{\lambda}, \qquad \Psi_{\text{quad}}(\theta) = \frac{\theta_+^2}{2\lambda}, \qquad \theta_+ := \max(\theta, 0). \tag{3.29}$$

We now apply this to (3.25) :

— In the branch H_0, the slope is $\theta = 1$, so $\sup_{\alpha \geq 0}\{\alpha - c(\alpha)\} = \Psi_{\text{quad}}(1) = \frac{1}{2\lambda}$. Hence

$$\sup_{\alpha \geq 0} H_0(p; \alpha) = \frac{1}{2\lambda} - \inf_{t \geq \max(0, p_{\max} - 1)} \left\{ \frac{\gamma}{2} \sum_{i=1}^{m} (p_i - t)_+^2 + \beta t \right\}. \tag{3.30}$$

— In the branch H_k, the slope is $\theta = p_k$, so

$$\sup_{\alpha \geq 0}\{\alpha p_k - c(\alpha)\} = \Psi_{\text{quad}}(p_k) = \frac{(p_k)_+^2}{2\lambda}.$$

Therefore

$$\sup_{\alpha \geq 0} H_k(p; \alpha) = \frac{(p_k)_+^2}{2\lambda} - \inf_{t \geq 0} \left\{ \frac{\gamma}{2} \Big((z_k^{(k)}(t))^2 + \sum_{i \neq k} (p_i - t)_+^2 \Big) + \beta t \right\}. \tag{3.31}$$

Finally, the quadratic-cost Hamiltonian is

$$H_{\text{quad}}(p) = \max \Big\{ \sup_{\alpha \geq 0} H_0(p; \alpha),\ \max_{1 \leq k \leq m} \sup_{\alpha \geq 0} H_k(p; \alpha) \Big\},$$

with (3.30) and (3.31).

Exponential instantaneous cost

We now solve for an exponential cost

$$c(\alpha) = \lambda_1 e^{\lambda_2 \alpha}, \qquad \lambda_1 > 0,\ \lambda_2 > 0.$$

For $\theta \in \mathbb{R}$, define the one-slope maximization problem

$$\Psi_{\text{exp}}(\theta) := \sup_{\alpha \geq 0} \Big\{ \alpha\, \theta - \lambda_1 e^{\lambda_2 \alpha} \Big\}. \tag{3.32}$$

Since the objective is strictly concave in α, the maximizer is characterized by the first-order condition

$$\theta - \lambda_1 \lambda_2 e^{\lambda_2 \alpha} = 0,$$

whenever the maximizer is interior. This yields the explicit optimizer/value :

$$\alpha_{\exp}(\theta) = \begin{cases} 0, & \theta \le \lambda_1\lambda_2, \\ \frac{1}{\lambda_2}\ln\Big(\frac{\theta}{\lambda_1\lambda_2}\Big), & \theta > \lambda_1\lambda_2, \end{cases} \qquad \Psi_{\exp}(\theta) = \begin{cases} -\lambda_1, & \theta \le \lambda_1\lambda_2, \\ \frac{\theta}{\lambda_2}\Big(\ln\frac{\theta}{\lambda_1\lambda_2} - 1\Big), & \theta > \lambda_1\lambda_2. \end{cases} \tag{3.33}$$

We now apply this to (3.25) :

— In the branch H_0, the slope is $\theta = 1$, so $\sup_{\alpha\ge0}\{\alpha - c(\alpha)\} = \Psi_{\exp}(1)$ and

$$\sup_{\alpha\ge0} H_0(p;\alpha) = \Psi_{\exp}(1) - \inf_{t\ge\max(0,p_{\max}-1)} \left\{ \frac{\gamma}{2}\sum_{i=1}^{m}(p_i - t)_+^2 + \beta t \right\}. \tag{3.34}$$

— In the branch H_k, the slope is $\theta = p_k$, hence

$$\sup_{\alpha\ge0} H_k(p;\alpha) = \Psi_{\exp}(p_k) - \inf_{t\ge0} \left\{ \frac{\gamma}{2}\Big((z_k^{(k)}(t))^2 + \sum_{i\ne k}(p_i - t)_+^2\Big) + \beta t \right\}. \tag{3.35}$$

Finally,

$$H_{\exp}(p) = \max\Big\{ \sup_{\alpha\ge0} H_0(p;\alpha),\ \max_{1\le k\le m}\sup_{\alpha\ge0} H_k(p;\alpha)\Big\},$$

with (3.34) and (3.35).

3.5.3 Resulting PDE and BSDE representation

The principal value u solves the reduced equation

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) + H(\nabla u(t,x)) = 0, \qquad u(T,x) = G(x),$$

where $H = H_{\text{quad}}$ in Subsection 3.5.2 and $H = H_{\exp}$ in Subsection 3.5.2.

We now also give a representation in terms of BSDEs. Let $(W_t)_{0\le t\le T}$ be an $\mathbb{R}^m$-Brownian motion and define the forward process

$$X_t := x_0 + W_t, \qquad 0 \le t \le T.$$

Let (Y, Z) be the solution of the BSDE

$$Y_t = G(X_T) + \int_t^T H(Z_s)\,ds - \int_t^T Z_s \cdot dW_s, \qquad 0 \le t \le T. \tag{3.36}$$

Then $u(0, x_0) = Y_0$. More generally, setting $u(t,x) := Y_t^{t,x}$ yields a (Markovian) viscosity solution of

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) + H(\nabla u(t,x)) = 0, \qquad u(T,x) = G(x).$$

This BSDE is solved numerically in the next section.

3.6 Numerics for the Principal–Agent game

We now illustrate the time development of the collapsed principal–agent problem on representative sample paths. Throughout this section we fix $m = 10$, $x_0 = 0_{\mathbb{R}^{10}}$, and $T = 1$. We take a quadratic instantaneous cost

$$c(\alpha) = \frac{1}{4}\alpha^2,$$

and the Blotto-type terminal gain

$$G(x) = S \sum_{i=1}^{m} w_i \, \mathbf{1}_{\{x_i \geq 0\}}, \qquad w_1 = 2, \quad w_i = 1 \ (i = 2, \ldots, m),$$

where $S > 0$ is a scaling (stakes) parameter. We solve the BSDE associated with the semilinear PDE

$$\partial_t u(t,x) + \tfrac{1}{2}\Delta u(t,x) + H(\nabla u(t,x)) = 0, \qquad u(T,x) = G(x),$$

using a neural-network Deep-BSDE scheme as described in Section 2.5.2. We use a Z-network with two hidden layers of size 500 and SiLU activation. We warm-start with a randomly initialized network and $y_0 = S \times \frac{\sum_i w_i}{2}$. After the warm start, we apply a generator relaxation as shown in 2.5.2 with $\delta_i = \frac{i}{5}$, $i = 1, 2, 3, 4, 5$. For each value δ_i, we train in three phases :

1. a first phase of 800 iterations with $N = 120$ time steps, $\lambda_{\text{res}} = 0.02$, and $\lambda_Z = 0.05$.
2. a second phase of 800 iterations with $N = 160$ time steps, $\lambda_{\text{res}} = 0.015$, and $\lambda_Z = 0.03$.
3. a final phase of 1000 iterations with $N = 200$ time steps, $\lambda_{\text{res}} = 0.01$, and $\lambda_Z = 0.02$.

Along a simulated Brownian trajectory $X_t = x_0 + W_t$, the Deep-BSDE solver returns processes $(Y_t, Z_t)_{0 \leq t \leq T}$, where Y_t approximates the Markovian value $u(t, X_t)$, and we have $Z_t \approx \nabla u(t, X_t)$. At each time t we reconstruct the principal's instantaneous optimal budget by maximizing the Isaacs Hamiltonian at the current gradient $p_t := Z_t$, i.e.

$$\alpha_t^\star = \alpha^\star(p_t),$$

using the closed-form rules derived in Section 3.5.2 and the branch decomposition $H = \max\{H_0, H_1\}$.

Remark 3.6.1. *A key qualitative implication of the Hamiltonian decomposition is the existence of two regimes. Under quadratic cost, whenever the no-investment branch H_0 dominates, the pointwise maximizer in α is constant : $\alpha_t^\star = 1/\lambda = 2$. Hence a plateau in $\alpha_t^\star$ is a structural signature of the H_0 regime. As the stakes parameter S increases, gradients become larger and the investment branch H_1 becomes optimal more frequently, producing state-dependent and time-varying budgets.*

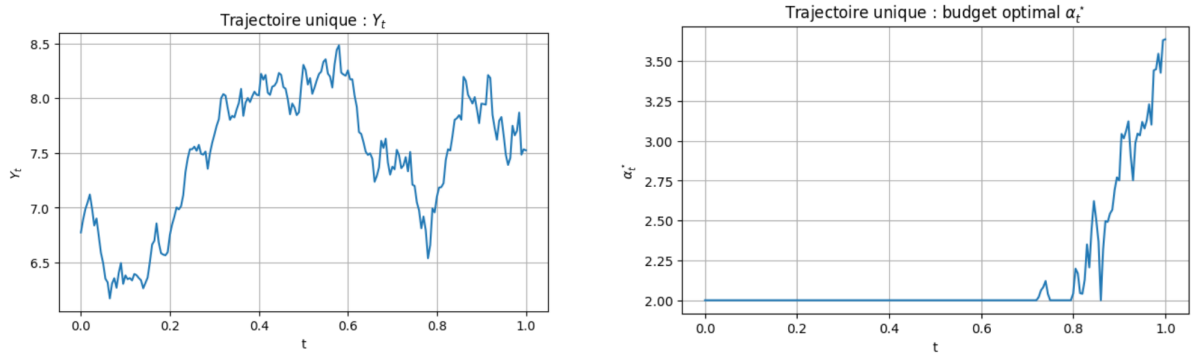


FIGURE 3.1 – Low-stakes regime ($S = 1$) : sample paths of Y_t and the reconstructed optimal budget $\alpha_t^\star$.

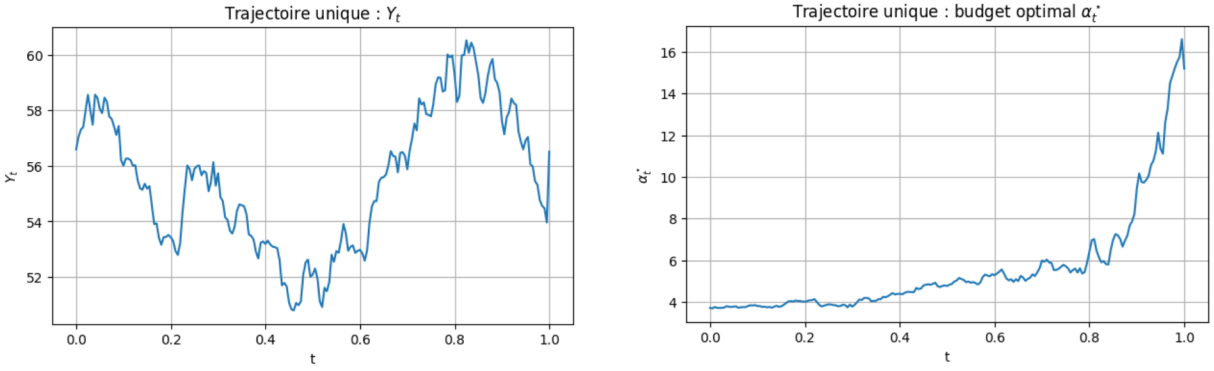


FIGURE 3.2 – High-stakes regime ($S = 8$) : sample paths of Y_t and the reconstructed optimal budget $\alpha_t^\star$.

Figures 3.1–3.2 illustrate the regime transition predicted by the Hamiltonian decomposition. Along each simulated Brownian trajectory, we reconstruct the principal's instantaneous optimal budget $\alpha_t^\star = \alpha^\star(p_t)$ by maximizing the pointwise Hamiltonian derived in Section 3.5.2.
In the low-stakes case ($S = 1$), the no-investment branch H_0 dominates over a large portion of the trajectory, which produces the characteristic plateau $\alpha_t^\star = 1/\lambda = 2$ under quadratic cost. Near the end of the trajectory, however, the optimizer may switch to an investment branch, leading to a visible departure from this plateau.
In the high-stakes case ($S = 8$), an investment branch dominates for most of the trajectory, leading to a strongly time-varying and state-dependent budget.
These observations are consistent with the theoretical structure : in the H_0 regime, the maximizer in α is constant under quadratic costs, whereas in an investment regime the optimal budget is driven by the effective slope in the Hamiltonian and becomes highly responsive to the realized path through $p_t = Z_t$.

Chapitre 4

The Mean-Field Blotto game

4.1 Introduction

Despite the richness of the generalizations of the Blotto game to asymmetric, stochastic, discrete, or multi-player contexts, the existing literature remains restricted to models with a finite number of players. However, studying a version with infinitely many players is naturally justified as soon as one seeks to model situations in which each agent faces a large crowd of opponents, so that the relevant interaction is with the aggregate distribution of strategies rather than with each individual player. For example, since the Blotto game is fundamentally a rank-based game, it fits particularly well into this framework, as one may wish to rank oneself relative to a crowd. To the best of the author's knowledge, no contribution has yet explored this approach within the framework of infinite population games or Mean Field Games (MFG).
The mean field approach, introduced independently by Lasry and Lions [53] and by Huang, Malhamé and Caines [46], consists in approximating strategic behavior in N-player symmetric games as N tends to infinity. Each player then interacts no longer with individualized opponents, but with an aggregate distribution representing the population. This makes it possible to transform a complex multi-player problem into an equation involving probability measures on the action space, linking the optimal individual strategy to the collective dynamics.
Our equilibrium concept is a static Cournot-Nash/Nonatomic equilibrium, closely related to mean-field interactions models. Existence results have been proven for various payoff functions with perfect or imperfect information, see for example Schmeidler [71] and Khan–Rustichini [2]. There is also a tight link with optimal transport problems as shown in the works of Blanchet and Carlier [9], [10].

In this chapter, we use "mean-field" in the large-population (non-atomic) sense : agents interact only through the population distribution. The model is static (one-shot), not dynamic.

Summary of the approach. Starting by defining the mean field Blotto formulation, we show that Dirac masses are trivial equilibria for it, then move to show that the problem is too general. An entropic penalization of Kullback–Leibler type is introduced in order to make the problem more specific. This modified problem is then proven to admit a unique best response of Gibbs form whose density is explicit as a function of the payoff. And the existence of an equilibrium is proven for a large set of infinite population, resource allocation problems. Then we move to the special case of the Blotto payoff, where the penalized best-reply problem is shown to admit

an explicit construcive equilibrium. In particular in the two-dimensional case, where an explicit equilibrium density on the simplex S_1 is derived.

Main contributions. In this chapter, a mean field version of the Blotto game is first introduced, in which players allocate continuous resources over m battlefields and interact only through the empirical distribution of allocations, thereby extending the classical finite-player Blotto literature to an infinite-population setting.
Then, the model is shown to admit both trivial Dirac equilibria and non-trivial diffuse equilibria under suitable choices of the rank-based payoff functions, highlighting a lack of structural uniqueness and suggesting directions for further refinements of the mean field Blotto framework.
An entropic Kullback–Leibler penalization is then incorporated into the individual optimization problem, and it is shown that this yields a unique best response for each population distribution, together with an existence result for mean field equilibria.
Finally, the structure of equilibria is analysed, with a particular focus on diffuse equilibria (absolutely continuous with respect to Lebesgue measure), and explicit formulas are obtained in the two-dimensional case $m = 2$ and three-dimensional case $m = 3$.

Structure of the chapter. Section 4.2 formally introduces the problem and the equilibrium we look to solve. Then, in Section 4.3, we show that the general problem needs more restriction through exhibiting trivial Dirac equilibria and other equilibria admitting a density w.r.t. the Lebesgue measure. Consequently, in Section 4.5, we introduce the chosen form of restruction, i.e. an etropy penalization. We introduce and motivate a modified version of the problem for a larger class of resource allocation rewards, and give general existence results for best responses and Mean-Field equilibria. Finally, in Section 4.6, we show that the special case of the Blotto game yields explicit equilibria and construct said equilibria in dimensions 2 and 3.

4.2 Model and assumptions

Consider a set of admissible actions $A \subset \mathbb{R}^m$, and a payoff function $J(a, \mu)$ representing the payoff of a player choosing strategy $a \in A$ against a population distribution $\mu \in \mathcal{P}(A)$, where $\mathcal{P}(A)$ denotes the space of probability measures on A. One can think of several classes of payoff functions $J : A \times \mathcal{P}(A) \to \mathbb{R}$ that fit the mean-field (nonatomic) setting. We list some examples :

Linear interaction through the mean. Let $\phi : A \to \mathbb{R}$ and $b : A \to \mathbb{R}^d$, and let $M(\mu) := \int_A b(y)\,\mu(dy) \in \mathbb{R}^d$. A simple class is

$$J(a, \mu) = \phi(a) + \langle c(a), M(\mu)\rangle,$$

where $c : A \to \mathbb{R}^d$. This covers, for instance, rewards that depend on how well a aligns with the population average.

Quadratic congestion / crowd-avoidance. Let $m(\mu) := \int_A y\,\mu(dy) \in \mathbb{R}^m$ be the mean action. A standard congestion cost is

$$J(a, \mu) = u(a) - \frac{\gamma}{2}\|a - m(\mu)\|^2, \qquad \gamma > 0,$$

or, conversely, a coordination incentive obtained by flipping the sign. Such payoffs depend only on the first moment of μ and are common in population games.

Pairwise interaction kernel. Let $K : A \times A \to \mathbb{R}$ be measurable and bounded. Define

$$J(a, \mu) = \int_A K(a, y)\, \mu(dy).$$

This is the static analogue of a McKean–Vlasov interaction : the payoff is the average interaction between a and the population.

Rank/quantile-based payoffs (Blotto-type). Assume $A = S_{m-1}$ is the simplex. Denote by μ_j the jth marginal of μ and by F_{μ_j} its CDF. Given increasing functions $g_j : [0, 1] \to \mathbb{R}$, define

$$J(a, \mu) = \sum_{j=1}^{m} g_j\big(F_{\mu_j}(a_j)\big).$$

A Blotto-type payoff is obtained by taking

$$J(a, \mu) = \sum_{j=1}^{m} \mathbf{1}_{\{F_{\mu_j}(a_j) \geq \alpha\}}, \qquad \alpha \in (0, 1),$$

meaning that one scores 1 on battlefield j whenever a_j lies in the top $(1 - \alpha)$ quantile of the population on that coordinate.

Threshold / knapsack-style rewards. Fix thresholds $q = (q_1, \ldots, q_m) \in [0, 1]^m$ and values $v = (v_1, \ldots, v_m) \in (0, \infty)^m$ and consider

$$J(a, \mu) = \sum_{j=1}^{m} v_j\, \mathbf{1}_{\{a_j \geq q_{\mu_j}(\alpha_j)\}},$$

with $q_{\mu_j}(\alpha_j)$ the α_j-quantile of μ_j. This makes explicit the endogenous-threshold structure behind rank-based games.

Spectral network payoff (non-additive across coordinates). Assume the coordinates of $a \in A \subset \mathbb{R}^m$ parametrize edge investments in a fixed graph $G = (V, E)$ with $|E| = m$. Let $w_j(a, \mu) = g_j(F_{\mu_j}(a_j))$ and let $L(a, \mu)$ be the associated weighted Laplacian. Define

$$J(a, \mu) = \lambda_2\big(L(a, \mu)\big),$$

the algebraic connectivity. This models a situation where local rank-based performance translates into a global network metric.

Risk-sensitive / robust payoff. Let $\ell(a, y)$ be a loss. For $\theta > 0$, a risk-sensitive criterion is

$$J(a, \mu) = -\frac{1}{\theta} \log\Big(\int_A e^{\theta \ell(a,y)}\, \mu(dy) \Big),$$

which interpolates between mean performance and worst-case behavior as θ increases.

Definition 4.2.1. *A* mean field equilibrium *is a measure* $\mu^* \in \mathcal{P}(A)$ *such that*

$$G(\mu^*) = \mu^*,$$

where

$$G(\nu) := \arg \max_{\mu \in \mathcal{P}(A)} \mathbb{E}_{U \sim \mu}\big[J(U, \nu)\big].$$

This means that μ^* is a measure that represents a population playing optimally against itself. Equivalently, μ^* can be defined as follows :

$$\mathrm{Supp}(\mu^*) \subset \arg\max_{a \in A} J(a, \mu^*),$$

that is, every point in the support of μ^* is an optimal strategy against the distribution μ^* itself.

A natural extension of the Blotto game to a mean field framework consists in considering m battlefields and an infinite population of players allocating resources across them. Let

$$S_{m-1} := \left\{ (x_1, \dots, x_m) \in [0,1]^m \;\middle|\; \sum_{i=1}^{m} x_i = 1 \right\}$$

denote the standard $(m-1)$-dimensional simplex, representing admissible budget allocation strategies. Let $\mathcal{L} := \mathcal{P}(S_{m-1})$ be the set of probability laws on this simplex.

Each player then faces the following optimization problem :

$$\sup_{\nu \in \mathcal{L}} \int_{S_{m-1}} \left[\sum_{i=1}^{m} g_i \left(F_{\mu_i}(x_i) \right) \right] d\nu(x), \tag{4.1}$$

where $\mu \in \mathcal{L}$ denotes the aggregate distribution of the population, and ν the strategic response of the player. Since the Blotto game is fundamentally rank-based, this formulation relies on an implicit comparison of each player with the entire population. This comparison is made possible through the introduction of the marginal distribution functions F_{μ_i} associated with each battlefield i, which measure the proportion of opponents allocating less resources to that battlefield.

The functions $g_i : [0,1] \to \mathbb{R}$ are assumed to be increasing, and encode the relative importance of achieving a favorable position on each battlefield. For instance, they may model the strategic or economic value attached to a high rank on battlefield i.

This framework makes it possible to study the existence of mean field equilibria in a context where players optimize their relative position in a dense population, rather than simply their absolute score in a two-player game.

However, there exist trivial equilibria : Dirac masses. Indeed, any Dirac mass clearly corresponds to the maximal payoff, since it induces a cumulative distribution function that is almost surely equal to 1.

This naturally raises the question of how to construct non-trivial equilibria for this game. In particular, one is interested in equilibria that are absolutely continuous with respect to Lebesgue measure, which will be referred to as **diffuse** equilibria.

4.3 A lack of structure

The previous result provides an example of a restricted framework in which one can find a diffuse equilibrium. However, when coming back to the original problem, one sees that the problem is in fact too general and therefore lacks structure. To see this, one can find, in certain situations, both trivial equilibria in the form of Dirac masses and other diffuse equilibria for the same problem.

For example, let π be a measure that is absolutely continuous with respect to Lebesgue measure. Consider the problem with $g_i \equiv F_{\pi_i}^{-1}$ for all $1 \leq i \leq m$, where $F_{\pi_i}^{-1}$ is the generalized inverse of the cumulative distribution function of π_i :

$$F_{\pi_i}^{-1}(x) := \inf\{y \in \mathbb{R} \mid F_{\pi_i}(y) \geq x\}.$$

For a population measure μ on S_{m-1}, recall that the best-response problem consists in maximizing

$$\nu \longmapsto \int_{S_{m-1}} \sum_{i=1}^{m} g_i\big(F_{\mu_i}(x_i)\big)\, d\nu(x), \tag{4.1}$$

and that a mean field equilibrium is a measure μ such that μ belongs to the set of best responses to μ.

In the present case, the choice $g_i \equiv F_{\pi_i}^{-1}$ completely freezes the structure of the game, while the population measure μ remains free. One first checks that π itself is a diffuse equilibrium for this game. Indeed, taking $\mu = \pi$, one obtains, for all $x \in S_{m-1}$,

$$g_i\big(F_{\pi_i}(x_i)\big) = F_{\pi_i}^{-1}\big(F_{\pi_i}(x_i)\big) = x_i,$$

where one uses that π_i is absolutely continuous (and, for instance, has a strictly positive density), so that F_{π_i} is continuous and strictly increasing, hence invertible. Consequently,

$$\sum_{i=1}^{m} g_i\big(F_{\pi_i}(x_i)\big) = \sum_{i=1}^{m} x_i = 1 \quad \text{for all } x \in S_{m-1},$$

and the best-response functional becomes

$$\int_{S_{m-1}} \sum_{i=1}^{m} g_i\big(F_{\pi_i}(x_i)\big)\, d\nu(x) = \int_{S_{m-1}} 1\, d\nu(x) = 1,$$

independently of the choice of ν. In other words, *every* measure ν is a best response to π, and in particular π is a diffuse mean field equilibrium, with equilibrium value equal to 1.

On the other hand, the same choice of functions $(g_i)_{1\leq i\leq m}$ also gives rise to purely atomic equilibria. Fix a point $x^\star = (x_1^\star, \dots, x_m^\star) \in S_{m-1}$ such that $x_i^\star > 0$ for all i, and consider the population measure $\mu = \delta_{x^\star}$. For each i, the marginal μ_i is then the Dirac mass at $x_i^\star$, and its cumulative distribution function is

$$F_{\mu_i}(y) = \begin{cases} 0, & y < x_i^\star, \\ 1, & y \geq x_i^\star. \end{cases}$$

It follows that, for all $x \in S_{m-1}$,

$$g_i\big(F_{\mu_i}(x_i)\big) = \begin{cases} F_{\pi_i}^{-1}(0), & x_i < x_i^\star, \\ F_{\pi_i}^{-1}(1), & x_i \geq x_i^\star. \end{cases}$$

In particular, at the point $x = x^\star$ one has $F_{\mu_i}(x_i^\star) = 1$ for all i, hence

$$\sum_{i=1}^{m} g_i\big(F_{\mu_i}(x_i^\star)\big) = \sum_{i=1}^{m} F_{\pi_i}^{-1}(1).$$

Since each term $g_i(F_{\mu_i}(x_i))$ is bounded above by $F_{\pi_i}^{-1}(1)$, the sum above is a maximum, attained only when $x = x^\star$ (because of the constraint $\sum_i x_i = 1$ and the positivity of the $x_i^\star$). Consequently, the best response to the population $\mu = \delta_{x^\star}$ is precisely to play $x^\star$, and $\delta_{x^\star}$ is therefore also a mean field equilibrium. The equilibrium value associated with $\delta_{x^\star}$ is

$$\int_{S_{m-1}} \sum_{i=1}^{m} g_i\big(F_{\mu_i}(x_i)\big)\, d\mu(x) = \sum_{i=1}^{m} F_{\pi_i}^{-1}(1),$$

which is typically strictly larger than 1. For example, if each marginal π_i is supported in $[0,1]$ with $F_{\pi_i}^{-1}(1) = 1$, one obtains an equilibrium value equal to m for $\delta_{x^\star}$, whereas the diffuse equilibrium π has value 1.

This example shows that with this choice of functions $(g_i)_{1\le i\le m}$ the same game simultaneously admits diffuse mean field equilibria and purely atomic equilibria, with different equilibrium values. In particular, there is no natural selection principle or Lasry–Lions type monotonicity property that would guarantee uniqueness (or even regularity) of the equilibrium. This illustrates that, in full generality, the problem under consideration is *too* flexible and does not possess the additional structure needed to obtain a well-defined diffuse equilibrium.

The next section introduces additional conditions in the form of a penalization, in order to study a slightly more structured version of the problem.

4.4 A toy mean-field model

Before turning to the model studied in this chapter, one can first look at a simplified "toy model". In this variant, one restricts attention to a class of probability measures that are absolutely continuous with respect to (normalized) Lebesgue measure on the simplex and whose densities satisfy additional regularity/compactness requirements. This restricted setting makes it possible to work with a compact strategy set and to obtain stronger regularity of the best-response map. Within this framework, one proves the existence of a mean-field equilibrium by combining these compactness properties with classical fixed-point arguments. The toy model and the corresponding results are presented in Appendix B.

4.5 Entropy-regularized mean-field allocation game

We elect to work with an entropic regularization. This choice is inspired by developments in optimal transport and in neural network optimization. In optimal transport, one optimizes a linear functional over a space of probability measures ; adding an entropic term makes the objective strictly convex (see [7]). Similarly, in neural network optimization, a non-convex problem over network parameters can be relaxed into a linear optimization over probability measures, which is then rendered strictly convex through an entropic regularization (see [18]). Such a penalization is therefore known to behave well, both theoretically and numerically.

Throughout this section, λ denotes the *normalized* Lebesgue measure on the simplex

$$S_{m-1} := \Big\{ x = (x_1, \dots, x_m) \in \mathbb{R}_+^m : \ \sum_{i=1}^{m} x_i = 1 \Big\}.$$

We write $\mathcal{L} \subset \mathcal{P}(S_{m-1})$ for the set of probability measures that are absolutely continuous with respect to λ.

Given $\nu \in \mathcal{P}(S_{m-1})$, denote by $f_\nu = \frac{d\nu}{d\lambda}$ its density when $\nu \ll \lambda$, and define the Kullback–Leibler divergence by

$$\mathrm{KL}(\nu, \lambda) = \begin{cases} \displaystyle\int_{S_{m-1}} f_\nu(x) \log\big(f_\nu(x)\big)\, d\lambda(x), & \text{if } \nu \ll \lambda, \\ +\infty, & \text{otherwise.} \end{cases}$$

Fix $\beta > 0$. For a general payoff function $R : S_{m-1} \times \mathcal{L} \to \mathbb{R}$ and a population $\mu \in \mathcal{L}$, the representative agent solves the entropy-regularized best-response problem

$$\sup_{\nu \in \mathcal{L}} \left\{ \int_{S_{m-1}} R(x, \mu)\, \nu(dx) \; - \; \frac{1}{\beta}\, \mathrm{KL}(\nu, \lambda) \right\}. \tag{4.1}$$

The criterion (4.1) is the standard form of a free-energy functional. The term $\int R\, d\nu$ plays the role of an average energy (or utility), while $\mathrm{KL}(\nu, \lambda)$ measures the cost of departing from the reference distribution. The parameter β is naturally interpreted as an inverse temperature : at high temperature (small β) the optimizer remains diffuse and close to λ, whereas at low temperature (large β) it concentrates on allocations with high payoff. In resource-allocation contexts, this interpretation is natural : it captures unmodelled fluctuations, limited precision, heterogeneity, or noise in decision-making through a single temperature parameter.

The term $\mathrm{KL}(\nu, \lambda)$ can also be read as a deformation cost of the uniform baseline. Moving from a neutral distribution to a distribution with sharp peaks or strong concentration on thin regions of the simplex is penalized. In particular, the fact that $\mathrm{KL}(\nu, \lambda) = +\infty$ as soon as ν is not absolutely continuous with respect to λ formalizes the idea that purely atomic (or more generally singular) strategies are idealized limits that are infinitely costly relative to a continuous baseline.

From an information-theoretic viewpoint, $\mathrm{KL}(\nu, \lambda)$ quantifies decision complexity : distorting the uniform distribution in a continuous action space amounts to injecting information (or computational effort) to discriminate among nearby allocations. One may equivalently view (4.1) as the Lagrangian form of a constrained problem : maximize $\int R\, d\nu$ under a complexity budget $\mathrm{KL}(\nu, \lambda) \le \kappa$, with β acting as the Lagrange multiplier. Varying β then parametrizes a continuum of regimes, from very diffuse to very concentrated strategies.

Finally, the entropy regularization is independent of the specific structure of R. Earlier in this manuscript, a rank-separable payoff of the form

$$R(x, \mu) = \sum_{i=1}^{m} g_i\big(F_{\mu_i}(x_i)\big)$$

was considered (indices correspond to marginals). This fits the present framework, but (4.1) also accommodates non-additive, spectral, or otherwise coupled rewards without changing the formalism. The entropy term thus serves as a common modeling block across a broader class of allocation games.

Another (formal) benefit is that the added term $-\frac{1}{\beta}\mathrm{KL}(\nu, \lambda)$ makes the objective strictly concave in ν, leading to a unique best response that can be written explicitly ; this is made precise in the next sections.

4.5.1 Entropy-regularized best response

For $\mu \in \mathcal{L}$, consider

$$\sup_{\nu \in \mathcal{L}} \left\{ \int_{S_{m-1}} R(x,\mu)\, \nu(dx) \; - \; \frac{1}{\beta}\, \mathrm{KL}(\nu, \lambda) \right\}. \tag{4.2}$$

Lemma 4.5.1. *Assume that $x \mapsto R(x,\mu)$ is bounded and measurable for every $\mu \in \mathcal{L}$. Then (4.2) admits a unique maximizer $\nu = \mathrm{BR}(\mu)$, given by*

$$\mathrm{BR}(\mu)(dx) = \frac{\exp(\beta R(x,\mu))}{Z(\mu)}\, \lambda(dx), \qquad Z(\mu) = \int_{S_{m-1}} \exp(\beta R(y,\mu))\, \lambda(dy). \tag{4.3}$$

In particular, $\mathrm{BR}(\mu) \ll \lambda$ with density $f_{\mathrm{BR}(\mu)}(x) = \exp(\beta R(x,\mu))/Z(\mu)$.

Proof. Let $\nu \ll \lambda$ with density $f = \frac{d\nu}{d\lambda}$. The objective in (4.2) becomes

$$\int_{S_{m-1}} R(x,\mu)\, f(x)\, d\lambda(x) \; - \; \frac{1}{\beta} \int_{S_{m-1}} f(x) \log f(x)\, d\lambda(x), \qquad f \geq 0, \; \int f\, d\lambda = 1.$$

Introduce a Lagrange multiplier η for the constraint $\int f\, d\lambda = 1$ and consider

$$\mathcal{J}(f) = \int R(x,\mu) f(x)\, d\lambda(x) - \frac{1}{\beta} \int f(x) \log f(x)\, d\lambda(x) - \eta\Big(\int f\, d\lambda - 1 \Big).$$

Pointwise optimality yields, for λ-a.e. x,

$$0 = \frac{\delta \mathcal{J}}{\delta f}(x) = R(x,\mu) - \frac{1}{\beta}\big(1 + \log f(x)\big) - \eta,$$

hence $\log f(x) = \beta(R(x,\mu) - \eta) - 1$ and therefore $f(x) \propto \exp(\beta R(x,\mu))$. Normalizing gives (4.3). Strict concavity of $f \mapsto -\int f \log f\, d\lambda$ on $\{f \geq 0, \; \int f\, d\lambda = 1\}$ implies uniqueness. □

Remark 4.5.2. *The best response is explicit and absolutely continuous with respect to λ ; in particular, optimal strategies cannot create atomic masses (which would have infinite KL cost).*

4.5.2 General mean-field equilibrium

Definition 4.5.3 (Best-response operator)**.** *Define the best-response operator*

$$T : \mathcal{P}(S_{m-1}) \to \mathcal{P}(S_{m-1}), \qquad T(\mu) := \mathrm{BR}(\mu),$$

where $\mathrm{BR}(\mu)$ is given by (4.3).

Definition 4.5.4 (Mean-field equilibrium)**.** *A* mean-field equilibrium *(MFE) is a fixed point $\mu^\star \in \mathcal{P}(S_{m-1})$ of T, i.e.*

$$\mu^\star = T(\mu^\star).$$

Assumption 4.5.5. *(A1) (Boundedness) There exists $C < \infty$ such that $|R(x,\mu)| \leq C$ for all $x \in S_{m-1}$ and all $\mu \in \mathcal{L}$.*

(A2) (Continuity in μ) If $\mu_n \Rightarrow \mu$ weakly, then $R(x,\mu_n) \to R(x,\mu)$ for λ-a.e. $x \in S_{m-1}$.

Theorem 4.5.6 (Existence of an MFE)**.** *Under Assumption 4.5.5, the map T is weakly continuous on $\mathcal{P}(S_{m-1})$ and admits at least one fixed point. In particular, an MFE exists.*

Proof. Since S_{m-1} is compact, $\mathcal{P}(S_{m-1})$ is compact and convex for the weak topology. Let $\mu_n \Rightarrow \mu$ and let $\varphi \in C(S_{m-1})$. By (4.3),

$$\int \varphi \, dT(\mu_n) = \frac{\int \varphi(x) \exp(\beta R(x, \mu_n)) \, \lambda(dx)}{\int \exp(\beta R(x, \mu_n)) \, \lambda(dx)}.$$

By (A1), $\exp(\beta R(\cdot, \mu_n)) \leq e^{\beta C}$. By (A2), $\exp(\beta R(x, \mu_n)) \to \exp(\beta R(x, \mu))$ for λ-a.e. x. Dominated convergence yields convergence of both numerator and denominator to the corresponding expressions with μ; the denominator stays strictly positive. Hence $\int \varphi \, dT(\mu_n) \to \int \varphi \, dT(\mu)$, i.e. $T(\mu_n) \Rightarrow T(\mu)$. Schauder–Tychonoff's fixed-point theorem (see Corollary 16.52 in [3]) then implies that T admits a fixed point. □

Remark 4.5.7. *The rank-separable reward*

$$R(x, \mu) = \sum_{i=1}^{m} g_i\big(F_{\mu_i}(x_i)\big)$$

fits the present framework : the best-response operator remains (4.3)*. Presenting the general case first makes it possible to vary R later (non-additive, spectral, complementarity effects, etc.) without changing either the best-response structure or the definition of equilibrium.*

4.6 Constructing a mean-field equilibrium for the Blotto game

We now focus on the Blotto-like specification

$$R_\alpha(x, \mu) = \sum_{i=1}^{m} \mathbf{1}_{\{F_{\mu_i}(x_i) \geq \alpha\}} = \sum_{i=1}^{m} \mathbf{1}_{\{x_i \geq q_{\mu_i}(\alpha)\}}, \qquad \alpha \in (0, 1), \tag{4.1}$$

where $q_{\mu_i}(\alpha) = F^{-1}_{\mu_i}(\alpha)$ denotes the α-quantile of the ith marginal. This quantile formulation is natural in Blotto because the game is fundamentally a *ranking* game : on each battlefield i, the agent's payoff depends on her position relative to the population, not on the absolute value of x_i. The condition $F_{\mu_i}(x_i) \geq \alpha$ means that at least an α-fraction of the population allocates no more than x_i on battlefield i, i.e. x_i exceeds the endogenous threshold $q_{\mu_i}(\alpha)$.

4.6.1 Homogeneous case and exchangeable equilibria

In the homogeneous case (where battlefields have the same value, see Subsection 1.2.3), it is natural to seek equilibria among laws invariant under permutations of coordinates.

Definition 4.6.1 (Exchangeable laws)**.** *A probability measure $\mu \in \mathcal{P}(S_{m-1})$ is* exchangeable *if, for every permutation σ of $\{1, \dots, m\}$, the image of μ under $x \mapsto (x_{\sigma(1)}, \dots, x_{\sigma(m)})$ equals μ. In that case, all marginals $\mu_1, \dots, \mu_m$ coincide ; we write $\mu_{(1)}$ for the common marginal and F_μ for its CDF.*

Definition 4.6.2 (One-parameter Gibbs family)**.** *For $q \in [0, 1]$, define $\mu_q \in \mathcal{P}(S_{m-1})$ by*

$$\mu_q(dx) = \frac{\exp\big(\beta \, s_q(x)\big)}{Z_m(q)} \, \lambda(dx), \qquad s_q(x) = \sum_{i=1}^{m} \mathbf{1}_{\{x_i \geq q\}}, \qquad Z_m(q) = \int_{S_{m-1}} \exp\big(\beta \, s_q(y)\big) \, \lambda(dy). \tag{4.2}$$

The measure μ_q is exchangeable ; denote by $F_{(q)}$ the CDF of its common marginal.

Proposition 4.6.3 (Quantile fixed point ⇒ exchangeable MFE)**.** *Fix $\alpha \in (0,1)$. If there exists $q^\star \in [0,1]$ such that*

$$F_{(q^\star)}(q^\star) = \alpha \qquad \textit{equivalently} \qquad q^\star = F^{-1}_{(q^\star)}(\alpha), \tag{4.3}$$

then $\mu_{q^\star}$ is a mean-field equilibrium for the Blotto payoff (4.1) within the exchangeable class.

Proof. Fix $q^\star$ satisfying (4.3) and consider the population $\mu_{q^\star}$. Since $\mu_{q^\star}$ is exchangeable, all marginals have CDF $F_{(q^\star)}$. The condition (4.3) says that the α-quantile of the common marginal equals $q^\star$, hence for every i,

$$\mathbf{1}_{\{F_{\mu_{q^\star,i}}(x_i)\geq\alpha\}} = \mathbf{1}_{\{x_i\geq q^\star\}}.$$

Therefore

$$R_\alpha(x,\mu_{q^\star}) = \sum_{i=1}^m \mathbf{1}_{\{x_i\geq q^\star\}} = s_{q^\star}(x).$$

By Lemma 4.5.1, the entropy-regularized best response to $\mu_{q^\star}$ has density proportional to $\exp(\beta R_\alpha(x,\mu_{q^\star})) = \exp(\beta s_{q^\star}(x))$ with respect to λ, which is precisely $\mu_{q^\star}$ by (4.2). Thus $\mathrm{BR}(\mu_{q^\star}) = \mu_{q^\star}$. □

Remark 4.6.4. *Proposition 4.6.3 reduces the search for an exchangeable MFE to a scalar fixed-point equation in q. This replaces a fixed-point problem on the infinite-dimensional space $\mathcal{P}(S_{m-1})$ by a one-dimensional equation.*

4.6.2 Vector thresholds and weighted battlefields

Definition 4.6.5 (Vector-threshold Gibbs family)**.** *Fix $\alpha = (\alpha_1,\dots,\alpha_m) \in (0,1)^m$ and $v = (v_1,\dots,v_m) \in (0,\infty)^m$. For $q = (q_1,\dots,q_m) \in [0,1]^m$, define $\mu_q^v \in \mathcal{P}(S_{m-1})$ by*

$$\mu_q^v(dx) := \frac{1}{Z_v(q)} \exp\Big(\beta\, s_q^v(x)\Big)\, \lambda(dx), \qquad s_q^v(x) := \sum_{i=1}^m v_i\, \mathbf{1}_{\{x_i\geq q_i\}},$$

$$Z_v(q) := \int_{S_{m-1}} \exp\Big(\beta\, s_q^v(y)\Big)\, \lambda(dy).$$

Denote by $F_i^{(v,q)}$ the CDF of the ith marginal of μ_q^v.

Proposition 4.6.6. *Consider the weighted Blotto payoff*

$$R_{\alpha,v}(x,\mu) := \sum_{i=1}^m v_i\, \mathbf{1}_{\{F_{\mu_i}(x_i)\geq\alpha_i\}}.$$

If $q^\star \in [0,1]^m$ satisfies $F_i^{(v,q^\star)}(q_i^\star) = \alpha_i$ for all $i = 1,\dots,m$, then $\mu_{q^\star}^v$ is an MFE for $R_{\alpha,v}$.

Proof. Fix such a $q^\star$ and consider $\mu_{q^\star}^v$. By definition of the quantiles, for each i,

$$\mathbf{1}_{\{F_{\mu^v_{q^\star,i}}(x_i)\geq\alpha_i\}} = \mathbf{1}_{\{x_i\geq q_i^\star\}},$$

hence

$$R_{\alpha,v}(x,\mu_{q^\star}^v) = \sum_{i=1}^m v_i\, \mathbf{1}_{\{x_i\geq q_i^\star\}} = s_{q^\star}^v(x).$$

Lemma 4.5.1 then implies that $\mathrm{BR}(\mu_{q^\star}^v)$ has density proportional to $\exp(\beta s_{q^\star}^v(x))$ with respect to λ, i.e. it equals $\mu_{q^\star}^v$ by Definition 4.6.5. □

Remark 4.6.7. *One does not need to prove existence of a suitable threshold vector $q^\star$ separately. Existence of an MFE $\mu^\star$ is already guaranteed by Theorem 4.5.6. From such an equilibrium, one can extract thresholds by setting $q_i^\star$ to be the α_i-quantile of the ith marginal of $\mu^\star$. These thresholds reproduce the ranking conditions in the payoff, and the Gibbs structure of the best response forces consistency.*

4.6.3 Explicit equilibria in low dimensions

Let us now give explicit equilibria for dimensions 2 and 3 in the Blotto game. We first discuss the two-dimensional case :

Dimension 2

Proposition 4.6.8. *Fix $\alpha \in (0,1)$ and set $a := e^\beta > 1$. At the exchangeable MFE, one draws $(X, 1-X) \in S_1$ with $X \in [0,1]$ and density (with respect to Lebesgue on $[0,1]$)*

$$f^\star(x) = \begin{cases} \dfrac{1+2\alpha(a-1)}{a}, & 0 \le x < \dfrac{a\alpha}{1+2\alpha(a-1)}, \\ 1+2\alpha(a-1), & \dfrac{a\alpha}{1+2\alpha(a-1)} \le x \le 1 - \dfrac{a\alpha}{1+2\alpha(a-1)}, \\ \dfrac{1+2\alpha(a-1)}{a}, & 1 - \dfrac{a\alpha}{1+2\alpha(a-1)} < x \le 1, \end{cases} \qquad \text{if } \alpha \le \frac{1}{2},$$

and

$$f^\star(x) = \begin{cases} a(2\alpha-1)+2(1-\alpha), & 0 \le x \le \dfrac{1-\alpha}{a(2\alpha-1)+2(1-\alpha)}, \\ \dfrac{a(2\alpha-1)+2(1-\alpha)}{a}, & \dfrac{1-\alpha}{a(2\alpha-1)+2(1-\alpha)} < x < \dfrac{a(2\alpha-1)+(1-\alpha)}{a(2\alpha-1)+2(1-\alpha)}, \\ a(2\alpha-1)+2(1-\alpha), & \dfrac{a(2\alpha-1)+(1-\alpha)}{a(2\alpha-1)+2(1-\alpha)} \le x \le 1, \end{cases} \qquad \text{otherwise.}$$

Proof. For $m = 2$, the density of μ_q in Definition 4.6.2 is piecewise constant on $[0,1]$ (depending on whether $x \ge q$ and/or $1-x \ge q$). One computes $F_{(q)}(q)$ explicitly in the two regimes $q \le \frac{1}{2}$ and $q > \frac{1}{2}$, solves $F_{(q)}(q) = \alpha$ for $q = q^\star$, and substitutes $q^\star$ back into the piecewise-constant Gibbs density. The resulting density is exactly the expression displayed above. □

Proposition 4.6.8 exhibits a two-phase structure, with a qualitative change at the critical value $\alpha = \frac{1}{2}$. We remind that $a = e^\beta > 1$. The equilibrium density is piecewise constant and is determined by the threshold

$$q^\star = q^\star(\alpha, \beta) = F^{-1}_{(q^\star)}(\alpha),$$

where $F_{(q)}$ denotes the common marginal CDF under μ_q from Definition 4.6.2. In dimension 2 the equilibrium decomposes into intervals on which one wins two battlefields, exactly one battlefield, or none, and the relevant regions are separated by $q^\star$ and $1-q^\star$.
The role of β is controling how sharply the equilibrium favors the payoff-maximizing regions, without changing which regions are preferred :

$$\frac{\text{high density}}{\text{low density}} \;=\; a \;=\; e^\beta,$$

The limit $\beta \to 0$ (strong regularization, $a \to 1$) collapses the two plateaux into a single one : $f^\star$ becomes uniform on $[0,1]$ and the threshold converges to the uniform quantile $q^\star \to \alpha$. At the opposite extreme, $\beta \to \infty$ (weak regularization, $a \to \infty$) forces concentration on allocations that maximize the number of wins subject to the ranking constraint imposed by α. If $\alpha < \frac{1}{2}$, winning both battlefields is feasible and the equilibrium concentrates near the balanced split $X = \frac{1}{2}$. If $\alpha > \frac{1}{2}$, winning both is impossible (since $q^\star > \frac{1}{2}$), and the equilibrium becomes bimodal, concentrating near the corners $X \approx 0$ and $X \approx 1$. Exactly at $\alpha = \frac{1}{2}$, the two-win region collapses to a point and the equilibrium is uniform for every β.
Varying α changes the ranking requirement : increasing α raises the threshold $q^\star$ and thus makes each win harder. For $\alpha \leq \frac{1}{2}$, the two-wins interval $[q^\star, 1-q^\star]$ shrinks as $\alpha \to \frac{1}{2}$, pushing the equilibrium toward near-balanced allocations. For $\alpha > \frac{1}{2}$, the no-win interval $(1-q^\star, q^\star)$ expands as α increases, while the density contrast between tails and middle remains governed by e^β.

Overall, α controls feasibility and the geometry of win regions, while β controls how concentrated the population is on the most rewarding ones.

Dimension 3

Definition 4.6.9. *Fix $\alpha \in (0,1)$ and set $a := e^\beta > 1$. Define*

$$\alpha_{1/3} := \frac{a+4}{3(a+2)}, \qquad \alpha_{1/2} := \frac{2a+1}{3a+1}.$$

Define $t = t(\alpha) \in (0,1)$ as the unique solution (in the interval indicated) of

$$\begin{cases} A_1 t^2 + B_1 t + C_1 = 0, & \text{if } \alpha \leq \alpha_{1/3} \quad (\text{with } t \in (0, \frac{1}{3}]), \\ A_2 t^2 + B_2 t + C_2 = 0, & \text{if } \alpha_{1/3} < \alpha \leq \alpha_{1/2} \quad (\text{with } t \in (\frac{1}{3}, \frac{1}{2}]), \\ A_3 t^2 + B_3 t + C_3 = 0, & \text{if } \alpha > \alpha_{1/2} \quad (\text{with } t \in (\frac{1}{2}, 1)), \end{cases}$$

where

$$A_1 = 4 - 5a + \alpha(15a - 9a^2 - 6), \quad B_1 = 2a + \alpha(6a^2 - 6a), \quad C_1 = -\alpha a^2,$$
$$A_2 = 4a^2 - 14a + 9 + \alpha(-12a^2 + 21a - 9),$$
$$B_2 = -4a^2 + 12a - 6 + \alpha(12a^2 - 18a + 6), \quad C_2 = (a-1)^2 + \alpha(-3a^2 + 3a - 1),$$
$$A_3 = 2a - 3 + \alpha(-3a + 3), \quad B_3 = -4a + 6 + \alpha(6a - 6), \quad C_3 = 2a - 2 + \alpha(-3a + 2).$$

Definition 4.6.10. *For $t \in (0,1)$ and $x = (x_1, x_2, x_3) \in S_2$, set*

$$k_t(x) := \#\{i \in \{1,2,3\} : x_i \geq t\}.$$

Define the density (with respect to Lebesgue on S_2)

$$f_t(x) := \frac{a^{k_t(x)}}{Z(t)},$$

where $Z(t)$ is given by

$$Z(t) = \begin{cases} aA_1(t) + a^2 A_2(t) + a^3 A_3(t), & t \leq \frac{1}{3}, \\ A_0(t) + aA_1(t) + a^2 A_2(t), & \frac{1}{3} < t \leq \frac{1}{2}, \\ A_0(t) + aA_1(t), & t > \frac{1}{2}, \end{cases}$$

with

$$\text{if } t \le \frac{1}{3}: \quad A_3(t) = \frac{(1-3t)^2}{2}, \quad A_2(t) = \frac{3\big((1-2t)^2 - (1-3t)^2\big)}{2}, \quad A_1(t) = \frac{1}{2} - A_2(t) - A_3(t),$$

$$\text{if } \frac{1}{3} < t \le \frac{1}{2}: \quad A_2(t) = \frac{3(1-2t)^2}{2}, \quad A_0(t) = \frac{(3t-1)^2}{2}, \quad A_1(t) = \frac{1}{2} - A_0(t) - A_2(t),$$

$$\text{if } t > \frac{1}{2}: \quad A_1(t) = \frac{3(1-t)^2}{2}, \quad A_0(t) = \frac{1}{2} - A_1(t).$$

Let μ_t denote the law on S_2 with density f_t.

Proposition 4.6.11 (Explicit exchangeable MFE in dimension 3)**.** *Fix $\alpha \in (0,1)$, let $t = t(\alpha)$ be as in Definition 4.6.9, and set $\mu^\star := \mu_{t(\alpha)}$ as in Definition 4.6.10. Then $\mu^\star$ is an exchangeable MFE for the homogeneous Blotto payoff*

$$R_\alpha(x,\mu) = \sum_{i=1}^{3} \mathbf{1}_{\{F_{\mu_i}(x_i) \ge \alpha\}}.$$

Equivalently, the equilibrium density is

$$f^\star(x) = \frac{a^{k_{t(\alpha)}(x)}}{Z(t(\alpha))}, \qquad x \in S_2.$$

Proof. For any threshold t, the law μ_t is exchangeable by construction. By Proposition 4.6.3 (with $m = 3$), it suffices to enforce the scalar fixed-point condition $F_{(t)}(t) = \alpha$, where $F_{(t)}$ is the common marginal CDF under μ_t. One computes $F_{(t)}(t)$ by splitting the simplex into regions according to how many coordinates exceed t. The geometry changes at $t = \frac{1}{3}$ and $t = \frac{1}{2}$, which produces a piecewise rational expression for $F_{(t)}(t)$. Writing $F_{(t)}(t) = \alpha$ in each regime reduces to a quadratic equation in t. The coefficients of these quadratics are precisely (A_r, B_r, C_r) in Definition 4.6.9, and the regime determines which root belongs to the corresponding interval. This defines $t(\alpha)$ and ensures $F_{(t(\alpha))}(t(\alpha)) = \alpha$. Proposition 4.6.3 then implies that $\mu_{t(\alpha)}$ is an exchangeable MFE. □

In dimension 3, Proposition 4.6.11 has the same basic structure as in dimension 2 : the equilibrium density is piecewise constant and favors allocations that win more battlefields by a multiplicative factor $a = e^\beta$ per additional win. However, the simplex geometry produces three distinct regimes, separated by the two critical levels $\alpha_{1/3}$ and $\alpha_{1/2}$ introduced in Definition 4.6.9.
The equilibrium is parametrized by the endogenous cutoff $t = t(\alpha)$ introduced in Definition 4.6.9 and the associated Gibbs law μ_t introduced in Definition 4.6.10. We remind :

$$k_t(x) = \#\{i \in \{1,2,3\} : \ x_i \ge t\}, \qquad f^\star(x) = \frac{a^{k_{t(\alpha)}(x)}}{Z(t(\alpha))},$$

one sees that moving from k to $k+1$ wins multiplies the local density by $a = e^\beta$. Thus, for a fixed threshold t, increasing β simply shifts probability mass toward regions with larger $k_t(x)$.
The distinction between regimes is governed by feasibility constraints on the simplex $x_1+x_2+x_3 = 1$. If $t \le \frac{1}{3}$, it is possible to have three coordinates simultaneously above t, so a triple-win region exists and contributes with weight a^3, which corresponds to the regime $\alpha \le \alpha_{1/3}$. If $\frac{1}{3} < t \le \frac{1}{2}$, triple wins are impossible but double wins remain feasible, so the best one can do is to win exactly two battlefields, which corresponds to the regime $\alpha_{1/3} < \alpha \le \alpha_{1/2}$. If $t > \frac{1}{2}$, even winning two

battlefields becomes impossible, and at most one coordinate can exceed t, so only single wins can occur (regime $\alpha > \alpha_{1/2}$). In this sense, α selects which win-levels are geometrically attainable through the induced cutoff $t(\alpha)$.
As in dimension 2, the high-temperature limit $\beta \to 0$ (so $a \to 1$) makes the density essentially uniform on S_2. Then $t(\alpha)$ is close to the α-quantile of the uniform marginal, and the phase boundaries in Definition 4.6.9 converge to the corresponding uniform values. Conversely, as $\beta \to \infty$ (so $a \to \infty$), the equilibrium becomes sharply concentrated on the regions with the *maximal feasible* number of wins : near the barycenter $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ when triple wins are feasible, near "two-large-one-small" profiles when only double wins are feasible, and near the vertices $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ when only single wins are feasible.

Finally, it is useful to separate what depends on the quantile constraint from what depends on β. At equilibrium, by construction of $t(\alpha)$ one has $F_{(t(\alpha))}(t(\alpha)) = \alpha$ (see Proposition 4.6.11 and its proof), which implies for each coordinate

$$\mathbb{P}(X_i \geq t(\alpha)) = 1 - \alpha.$$

Hence the expected number of wins equals $3(1 - \alpha)$, independently of β. What β changes is the distribution of the number of wins (how often one wins exactly $0, 1, 2$, or 3 battlefields), through the relative weights $a^{k_t(x)}$ in Definition 4.6.10.

Remark 4.6.12. *If one was to study the equilibria for any dimension m, it would have similar characteristics as the ones we just showed : a plateau structure, distributed over multiple regions that depend on the parameter α. Thresholds would be observed at values $\alpha_i = \frac{1}{i}$, which define how many battlefields can be won. And β would play again the role of controlling the distribution of the number of wins.*

Annexe A

Preuves

This Section is devoted to the proof of Lemma 2.3.8.

Proof. Let $(t, x_1, x_2) \in [0,T) \times \mathbb{R}^m \times \mathbb{R}^m$. We first show that

$$w(t, x_1 - x_2) = \sup_{a \in \mathcal{A}} J(t, x_1, x_2, a, b^\star) = J(t, x_1, x_2, a^\star, b^\star)$$

Let $\delta \in (0, T-t)$ and set $\gamma := x_1 - x_2$. Writing $\Gamma^{t,\gamma} := X^{1,t,x_1} - X^{2,t,x_2}$ and using the smoothness of w, we may apply Itô's formula and obtain

$$\begin{aligned}
&\mathbb{E}^{P^{a,b^\star}}\left[w\left(T-\delta, X^{1,t,x_1}_{T-\delta} - X^{2,t,x_2}_{T-\delta}\right)\right] - w(t,\gamma) \\
&= \mathbb{E}^{P^{a,b^\star}}\left[\int_t^{T-\delta} \left[\partial_t w\left(s, \Gamma^{t,\gamma}_s\right) + \Delta w\left(s, \Gamma^{t,\gamma}_s\right) + (a - b^\star)\left(s, X^{1,t,x_1}_s, X^{2,t,x_2}_s\right).Dw\left(s, \Gamma^{t,\gamma}_s\right)\right] ds\right] \\
&\leq \mathbb{E}^{P^{a,b^\star}}\left[\int_t^{T-\delta} \left[\partial_t w(s, \Gamma^{t,\gamma}_s) + \Delta w(s, \Gamma^{t,\gamma}_s) - b^\star\left(s, X^{1,t,x_1}_s, X^{2,t,x_2}_s\right).Dw(s, \Gamma^{t,\gamma}_s) + \sup_{a \in A} a.Dw(s, \Gamma^{t,\gamma}_s)\right] ds\right] \\
&= 0,
\end{aligned}$$

using Lemma 2.3.7. The next step is to pass to the limit $\delta \to 0$ using the dominated convergence theorem. Let $(t,z) \in [0,T) \times \mathbb{R}^m$, we have using Definition 2.3.1 of w :

$$\begin{aligned}
|w(t,z)| &\leq \frac{1}{(4\pi(T-t))^{\frac{m}{2}}} \int_{\mathbb{R}^m} \exp\left[-\frac{|z-r|^2}{4(T-t)}\right] |g(r)| dr \\
&\leq \frac{1}{(4\pi(T-t))^{\frac{m}{2}}} \int_{\mathbb{R}^m} \exp\left[-\frac{|r|^2}{4(T-t)}\right] |g(r-z)| \, dr \\
&\leq C_1 \frac{1}{(4\pi(T-t))^{\frac{m}{2}}} \int_{\mathbb{R}^m} \exp\left[-\frac{|r|^2}{4(T-t)} + C_2 |r-z|\right] dr \\
&\leq C_1 \exp\left[C_2 |z|\right] \frac{1}{(4\pi(T-t))^{\frac{m}{2}}} \int_{\mathbb{R}^m} \exp\left[-\frac{|r|^2}{4(T-t)} + C_2 |r|\right] dr \\
&:= Q(t) C_1 \exp\left[C_2 |z|\right]
\end{aligned}$$

Where C_1 and C_2 are the constants from assumption 2.2.1. Let us now prove that $Q(t)$ can be bounded independently of t.

By equivalence of norms in finite dimension, there exist constants $B_1, B_2 > 0$ such that

$|r| \leq B_1 \sum_{i=1}^m |r_i|$ and $\sqrt{\sum_{i=1}^m |r_i|^2} \leq \frac{1}{B_2}|r|$ for all $r \in \mathbb{R}^m$. We have, using these notations

$$\begin{aligned}
Q(t) &= \frac{1}{(4\pi(T-t))^{\frac{m}{2}}} \int_{\mathbb{R}^m} \exp\left[-\frac{|r|^2}{4(T-t)} + C_2|r|\right] dr \\
&\leq \frac{1}{(4\pi(T-t))^{\frac{m}{2}}} \int_{\mathbb{R}^m} \exp\left[-B_2^2\frac{\sum_{i=1}^m |r_i|^2}{4(T-t)} + C_2B_1\sum_{i=1}^m |r_i|\right] dr \\
&= \prod_{i=1}^m \frac{1}{\sqrt{4\pi(T-t)}} \int_{\mathbb{R}} \exp\left[-B_2^2\frac{|r_i|^2}{4(T-t)} + C_2B_1|r_i|\right] dr_i \\
&= \left(\frac{1}{\sqrt{4\pi(T-t)}} \int_{\mathbb{R}} \exp\left[-B_2^2\frac{|r|^2}{4(T-t)} + C_2B_1|r|\right] dr\right)^m \\
&= \left(\frac{1}{\sqrt{4\pi(T-t)}} \left(\int_{\mathbb{R}_+} \exp\left[-B_2^2\frac{r^2}{4(T-t)} + C_2B_1 r\right] dr + \int_{\mathbb{R}_-} \exp\left[-B_2^2\frac{r^2}{4(T-t)} - C_2B_1 r\right] dr\right)\right)^m \\
&\leq \big(Q_1(t) + Q_2(t)\big)^m,
\end{aligned}$$

Where

$$\begin{aligned}
Q_1(t) &:= \frac{1}{\sqrt{4\pi(T-t)}} \int_{\mathbb{R}} \exp\left[-B_2^2\frac{r^2}{4(T-t)} + C_2B_1 r\right] dr \\
Q_2(t) &:= \frac{1}{\sqrt{4\pi(T-t)}} \int_{\mathbb{R}} \exp\left[-B_2^2\frac{r^2}{4(T-t)} - C_2B_1 r\right] dr
\end{aligned}$$

These are Gaussian integrals that can be bounded as follows :

$$\begin{aligned}
Q_1(t) &= \frac{1}{\sqrt{4\pi(T-t)}} \int_{\mathbb{R}} \exp\left[-B_2^2\frac{r^2}{4(T-t)} + C_2B_1 r\right] dr \\
&= \frac{1}{\sqrt{4\pi(T-t)}} \int_{\mathbb{R}} \exp\left[-\frac{B_2^2}{4(T-t)}\left(r - \frac{2C_2B_1(T-t)}{B_2^2}\right)^2 + \left(\frac{2C_2B_1(T-t)}{B_2^2}\right)^2\right] dr \\
&= \frac{\exp\left[\left(\frac{2C_2B_1(T-t)}{B_2^2}\right)^2\right]}{B_2} \\
&\leq \frac{\exp\left[\left(\frac{2C_2B_1T}{B_2^2}\right)^2\right]}{B_2}
\end{aligned}$$

$Q_2(t)$ can be bounded by the same constant in a similar manner.

Denoting $C_3 := \frac{C_1}{B_2^m} C_1 \left(2\exp\left[\left(\frac{2C_2B_1T}{B_2^2}\right)^2\right]\right)^m$, we finally have :

$$|w(t,z)| \leq C_3 \exp\left[C_2|z|\right]$$

Using the last inequality :

$$\begin{aligned}
\left|w\left(T-\delta, \Gamma_{T-\delta}^{t,\gamma}\right)\right| &\leq C_3 \exp\left[C_2\left|\Gamma_{T-\delta}^{t,\gamma}\right|\right] \\
&= C_3 \exp\left[C_2\left|x + \int_t^{T-\delta} (a-b^\star)\left(u, X_u^{1,t,x_1}, X_u^{2,t,x_2}\right) du + \int_t^{T-\delta} dW_u^1 - dW_u^2\right|\right] \\
&\leq C_3 \exp\left[C_2|x|\right] \exp\left[\left|\int_t^{T-\delta} (a-b^\star)\left(u, X_u^{1,t,x_1}, X_u^{2,t,x_2}\right) du\right|\right] \exp\left[\left|\int_t^{T-\delta} dW_u^1 - dW_u^2\right|\right]
\end{aligned} \tag{A.1}$$

We now aim to bound each term of the right hand side independently of δ to use dominated convergence.
We have $a, b^\star \in \mathcal{A}$ so they are bounded. Let $C_\mathcal{A}$ be the upper bound of $\mathcal{A}$, i.e. $|a(t,x_1,x_2)| \leq C_\mathcal{A}$ for all $a \in \mathcal{A}$ and $(t,x_1,x_2) \in [0,T) \times \mathbb{R}^m \times \mathbb{R}^m$. We have :

$$\begin{aligned}\exp\left[\left|\int_t^{T-\delta} (a-b^\star)\left(u, X_u^{1,t,x_1}, X_u^{2,t,x_2}\right) du\right|\right] &\leq \exp\left[\int_0^T \left|(a-b^\star)\left(u, X_u^{1,t,x_1}, X_u^{2,t,x_2}\right)\right| du\right] \\ &\leq \exp\left[2TC_\mathcal{A}\right]\end{aligned}$$

Plugging this in inequality (A.1) :

$$\begin{aligned}\left|w\left(T-\delta, \Gamma_{T-\delta}^{t,\gamma}\right)\right| &\leq C_3 \exp\left[C_2|\gamma| + 2TC_\mathcal{A}\right] \exp\left[\left|\int_t^{T-\delta} dW_u^1 - dW_u^2\right|\right] \\ &\leq C_4 \exp\left[\left|W_{T-\delta}^1 - W_t^1 - W_{T-\delta}^2 + W_t^2\right|\right]\end{aligned}$$

where $C_4 := C_3 \exp\left[C_2|\gamma| + 2TC_\mathcal{A}\right]$.
Finally, we know that W^1 and W^2 are $P^{a,b^\star}$-standard Brownian motions. We denote $W^i = (W^{i,j})_{j=1,\dots,m}$ the decomposition of W^i into m independent $P^{a,b^\star}$-standard Brownian motions on $\mathbb{R}$. We have :

$$\begin{aligned}\exp\left[\left|W_{T-\delta}^1 - W_t^1 - W_{T-\delta}^2 + W_t^2\right|\right] &\leq \exp\left[2 \sup_{u\in[0,T]} \left|W_u^1\right|\right] \exp\left[2 \sup_{u\in[0,T]} \left|W_u^2\right|\right] \\ &\leq \exp\left[2B_1\left(\sup_{u\in[0,T]} \sum_{i=1}^m \left|W_u^{1,i}\right|\right)\right] \exp\left[2B_1\left(\sup_{u\in[0,T]} \sum_{i=1}^m \left|W_u^{2,i}\right|\right)\right] \\ &\leq \exp\left[2B_1\left(\sum_{i=1}^m \sup_{u\in[0,T]} \left|W_u^{1,i}\right|\right)\right] \exp\left[2B_1\left(\sum_{i=1}^m \sup_{u\in[0,T]} \left|W_u^{2,i}\right|\right)\right] \\ &\leq \exp\left[2B_1\left(\sum_{i=1}^m \sup_{u\in[0,T]} W_u^{1,i} + \sup_{u\in[0,T]} -W_u^{1,i}\right)\right] \\ &\quad \times \exp\left[2B_1\left(\sum_{i=1}^m \sup_{u\in[0,T]} W_u^{2,i} + \sup_{u\in[0,T]} -W_u^{2,i}\right)\right] \\ &= \prod_{i=1}^m \exp\left[2B_1\left(\sup_{u\in[0,T]} W_u^{1,i} + \sup_{u\in[0,T]} -W_u^{1,i}\right)\right] \\ &\quad \times \prod_{i=1}^m \exp\left[2B_1\left(\sup_{u\in[0,T]} W_u^{2,i} + \sup_{u\in[0,T]} -W_u^{2,i}\right)\right]\end{aligned}$$

The expectation under $P^{a,b^\star}$ then yields :

$$\mathbb{E}^{P^{a,b^\star}}\left[\exp\left[\left|W_{T-\delta}^1 - W_t^1 - W_{T-\delta}^2 + W_t^2\right|\right]\right]$$

$$
\begin{aligned}
&\leq \mathbb{E}^{P^{a,b^\star}}\left[\prod_{i=1}^{m}\exp\left[2B_1\left(\sup_{u\in[0,T]}W_u^{1,i}+\sup_{u\in[0,T]}-W_u^{1,i}\right)\right]\prod_{i=1}^{m}\exp\left[2B_1\left(\sup_{u\in[0,T]}W_u^{2,i}+\sup_{u\in[0,T]}-W_u^{2,i}\right)\right]\right]\\
&= \mathbb{E}^{P^{a,b^\star}}\left[\prod_{i=1}^{m}\exp\left[2B_1\left(\sup_{u\in[0,T]}W_u^{1,i}+\sup_{u\in[0,T]}-W_u^{1,i}\right)\right]\right]\\
&\times \mathbb{E}^{P^{a,b^\star}}\left[\prod_{i=1}^{m}\exp\left[2B_1\left(\sup_{u\in[0,T]}W_u^{2,i}+\sup_{u\in[0,T]}-W_u^{2,i}\right)\right]\right]\\
&\leq \left(\mathbb{E}^{P^{a,b^\star}}\left[\prod_{i=1}^{m}\exp\left[4B_1\sup_{u\in[0,T]}W_u^{1,i}\right]\right]\mathbb{E}^{P^{a,b^\star}}\left[\prod_{i=1}^{m}\exp\left[4B_1\sup_{u\in[0,T]}-W_u^{1,i}\right]\right]\right)^{\frac{1}{2}}\\
&\times \left(\mathbb{E}^{P^{a,b^\star}}\left[\prod_{i=1}^{m}\exp\left[4B_1\sup_{u\in[0,T]}W_u^{2,i}\right]\right]\mathbb{E}^{P^{a,b^\star}}\left[\prod_{i=1}^{m}\exp\left[4B_1\sup_{u\in[0,T]}-W_u^{2,i}\right]\right]\right)^{\frac{1}{2}}\\
&= \left(\mathbb{E}^{P^{a,b^\star}}\left[\exp\left[4B_1\sup_{u\in[0,T]}W_u^{1,1}\right]\right]\right)^{2m} < \infty,
\end{aligned}
$$

where we used the independence between W^1 and W^2 and the Cauchy-Schwarz inequality. This means that $\left|w\left(T-\delta,\Gamma_{T-\delta}^{t,\gamma}\right)\right|$ is dominated by an integrable function that does not depend on δ. The limit of its expectation under $P^{a,b^\star}$ when $\delta\to 0$ within the inequality A.1 gives :

$$
\begin{aligned}
\mathbb{E}^{P^{a,b^\star}}\left[w\left(T-\delta,\Gamma_{T-\delta}^{t,\gamma}\right)\right]\leq w(t,\gamma) &\implies \lim_{\delta\to 0}\mathbb{E}^{P^{a,b^\star}}\left[w\left(T-\delta,\Gamma_{T-\delta}^{t,\gamma}\right)\right]\leq w(t,\gamma)\\
&\implies \mathbb{E}^{P^{a,b^\star}}\left[\lim_{\delta\to 0}w\left(T-\delta,\Gamma_{T-\delta}^{t,\gamma}\right)\right]\leq w(t,\gamma)\\
&\implies \mathbb{E}^{P^{a,b^\star}}\left[\mathbf{1}_{\mathbb{R}^m\setminus K}\left(\Gamma_T^{t,\gamma}\right)g\left(\Gamma_T^{t,\gamma}\right)\right]\leq w(t,\gamma)\\
&\implies \mathbb{E}^{P^{a,b^\star}}\left[g\left(\Gamma_T^{t,\gamma}\right)\right]\leq w(t,\gamma)\\
&\implies J(t,x_1,x_2,a,b^\star)\leq w(t,\gamma)
\end{aligned}
$$

Where K is the set of discontinuity of g. Since $\Gamma_T^{t,\gamma}$ follows a normal distribution, we have $P^{a,b^\star}\left(\Gamma_T^{t,\gamma}\in K\right)=0$. This is true for all $a\in\mathcal{A}$. So we have :

$$\sup_{a\in\mathcal{A}}J(t,x_1,x_2,a,b^\star)\leq w(t,\gamma)$$

The first Ito application in this proof gives us for $a=a^\star$:

$$\mathbb{E}^{P^{a,b^\star}}\left[w\left(T-\delta,X_{T-\delta}^{t,x}\right)\right]-w(t,x)=0$$

meaning $w(t,\gamma)=J(t,x_1,x_2,a^\star,b^\star)$ by taking $\delta\to 0$. We finally get :

$$J(t,x_1,x_2,a^\star,b^\star)\leq\sup_{a\in\mathcal{A}}J(t,x_1,x_2,a,b^\star)\leq w(t,\gamma)=J(t,x_1,x_2,a^\star,b^\star)$$

Which proves the equality. The proof that $\inf_{b\in\mathcal{A}}J(t,x_1,x_2,a^\star,b)=w(t,\gamma)$ is done in a similar manner. □

Annexe B

Modèle-jouet à champ moyen

One can state an existence result of a Mean Field equilibrium for a toy model in which the optimization domain is strongly restricted so as to recover compactness.

Assume that the $g_i : [0,1] \to \mathbb{R}^+$ form a family of m q-Lipschitz functions with $q > 0$, and let $c = (c_i)_{i \geq 0}$ be a sequence of positive real numbers.

Definition B.0.1. *Define the set $E_m^c \subset \mathcal{L}$ of measures ν that admit a density f_ν with respect to Lebesgue measure, such that*

$$\forall j \in \mathbb{N},\ \forall \alpha_j \in \mathbb{N}^m \text{ with } \sum_{i=1}^{m} \alpha_i = j,\ \forall x \in [0,1]^m, \quad \left| f_\nu^{\alpha_j}(x) \right| = \left| \frac{\partial^j}{\partial x_1^{\alpha_1} \cdots \partial x_m^{\alpha_m}} f_\nu(x) \right| \leq c_j.$$

In this setting, one uses Kakutani's fixed point theorem [48], which provides a fixed point result for set-valued maps :

Theorem B.0.2 (Kakutani fixed point theorem)**.** *Let K be a nonempty convex subset of a locally convex, Hausdorff topological vector space.*

Suppose that a map $\psi : K \to 2^K$ satisfies the following properties :

1. *For every $x \in K$, the set $\psi(x)$ is nonempty and convex.*
2. *The graph $\{(x,y) : x \in K,\ y \in \psi(x)\}$ is closed.*

Then ψ has a fixed point. In other words, there exists $x \in K$ such that $x \in \psi(x)$.

The following result of Berge [8] will also be needed :

Theorem B.0.3 (Berge's maximum theorem)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces, with $\mathcal{Y}$ compact. Let $F : \mathcal{X} \to 2^{\mathcal{Y}}$ and $g : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be given. Assume that g is jointly continuous and that F satisfies the following two properties :*

1. *The graph $Gr(F) := \{(x,y) : x \in \mathcal{X},\ y \in F(x)\}$ is closed.*
2. *F is lower hemicontinuous, that is : if $x_n \to x$ in $\mathcal{X}$ and $y \in F(x)$, then there exists a subsequence $n_1 < n_2 < \dots$ and points $y_{n_k} \in F(x_{n_k})$ such that $y_{n_k} \to y$.*

Define $v : \mathcal{X} \to \mathbb{R}$ and $\psi : \mathcal{X} \to 2^{\mathcal{Y}}$ by

$$v(x) = \sup_{y \in F(x)} g(x,y), \quad \psi(x) = \arg\max_{y \in F(x)} g(x,y).$$

Then v is continuous, $\psi(x) \neq \varnothing$ for all $x \in \mathcal{X}$, and the graph $Gr(\psi) := \{(x,y) : x \in \mathcal{X},\ y \in \psi(x)\}$ is closed.

This theorem is now applied to the map that associates, to a mean field type distribution, the set of optimal responses. This defines the correspondence $\Psi(\mu)$.

Berge's maximum theorem implies that, for every μ, the set $\Psi(\mu)$ is nonempty and that the graph of Ψ is closed. The convexity of $\Psi(\mu)$ then follows from the linearity of the function g defined below.

It remains to check that the assumptions of Berge's theorem are satisfied. In the present case, the function g is defined for all $\mu \in \mathcal{P}([0,1]^m)$ and $\nu \in E_m^c$ as follows.

Lemma B.0.4. *For a given sequence $c \in \mathbb{R}_+^{\mathbb{N}}$ and a given $m \in \mathbb{N}$, the set E_m^c is compact.*

Proof. Let $(\mu_n)_{n\geq 0}$ be a sequence in E_m^c.
We want to extract a subsequence that converges in E_m^c. The idea is to use the densities and the Arzela-Ascoli theorem a countably infinite number of times and use Cantor's diagonal extraction to find a limit. We denote by f_n the density of μ_n. Since all f_n satisfy property (C), (f_n) and all their derivatives are uniformly bounded.
We first prove that $f_n^{\alpha_j}$ is c_{j+1}-Lipschitz for any given order of derivation $j \in \mathbb{N}$ and vector $\alpha_j \in \mathbb{N}^m$ such that $\sum_{i=1}^m \alpha_i = j$.
For x and y in $[0,1]^m$:

$$\begin{aligned}
\left|f_n^{\alpha_j}(x) - f_n^{\alpha_j}(y)\right| &= \left|f_n^{\alpha_j}(x_1,\dots,x_{m-1},x_m) - f_n^{\alpha_j}(x_1,\dots,x_{m-1},y_m)\right.\\
&\qquad \left. + f_n^{\alpha_j}(x_1,\dots,x_{m-1},y_m) - f_n^{\alpha_j}(y_1,\dots,y_m)\right|\\
&\leq \left|f_n^{\alpha_j}(x_1,\dots,x_{m-1},x_m) - f_n^{\alpha_j}(x_1,\dots,x_{m-1},y_m)\right|\\
&\qquad + \left|f_n^{\alpha_j}(x_1,\dots,x_{m-1},y_m) - f_n^{\alpha_j}(y_1,\dots,y_m)\right|\\
&= \left|\int_{y_m}^{x_m} \frac{\partial^{j+1} f_n}{\partial x_1^{\alpha_1}\cdots\partial x_{m-1}^{\alpha_{m-1}}\partial x_m^{\alpha_m+1}}(x_1,\dots,x_{m-1},s)\,ds\right|\\
&\qquad + \left|f_n^{\alpha_j}(x_1,\dots,x_{m-1},y_m) - f_n^{\alpha_j}(y_1,\dots,y_m)\right|\\
&\leq c_{j+1}\left|x_m - y_m\right| + \left|f_n^{\alpha_j}(x_1,\dots,x_{m-1},y_m) - f_n^{\alpha_j}(y_1,\dots,y_m)\right|.
\end{aligned}$$

We use the same method for dimensions $m-1$ up to the first dimension to prove that $f_n^{\alpha_j}$ is c_{j+1}-Lipschitz.
This means for any derivation vector α_j, $\left(f_n^{\alpha_j}\right)_{n\geq 0}$ is a uniformly bounded sequence by c_j and uniformly equicontinuous (since they are all c_{j+1}-Lipschitz). So we can extract, through the Arzela-Ascoli theorem a uniformly converging subsequence for any α_j.
We conclude through a diagonal extraction argument that we can extract a subsequence $\left(f_{\phi(n)}\right)_{n\geq 0}$ of densities which converge uniformly (as well their derivatives) to a density f. f is then $\mathcal{C}^\infty$ and its derivatives are bounded by the values of c.
Let μ be the distribution defined by the limit density f. Then $\left(\mu_{\phi(n)}\right)_{n\geq 0}$ converge in distribution to μ by Scheffé's lemma. □

Lemma B.0.5. *Let $c \in \mathbb{R}_+^{\mathbb{N}}$ be a given sequence. Let g be the function defined for all $(\mu,\nu) \in \mathcal{L} \times E_m^c$ by*

$$g(\mu,\nu) := \int_{S_{m-1}} \left[\sum_{i=1}^m g_i\left(F_{\mu_i}(x_i)\right)\right] d\nu(x).$$

Then g is continuous with respect to the topology of convergence in law.

Proof. Let $(\mu^n)_{n\geq 0}$ (resp. $(\nu^n)_{n\geq 0}$) be a sequence in E_m^c converging to μ (resp. ν).
Let's prove that $g(\mu^n,\nu^n) \to g(\mu,\nu)$. We have :

$$
\begin{aligned}
\left|g(\mu^n,\nu^n)-g(\mu,\nu)\right| &= \left|g(\mu^n,\nu^n)-g(\mu,\nu^n)+g(\mu,\nu^n)-g(\mu,\nu)\right| \\
&\le \left|g(\mu^n,\nu^n)-g(\mu,\nu^n)\right| + \left|g(\mu,\nu^n)-g(\mu,\nu)\right| \\
&= \left|\int_{[0,1]^m}\sum_{i=1}^m \Big(g_i(F_{\mu_i^n}(x_i))-g_i(F_{\mu_i}(x_i))\Big)\,d\nu^n(x)\right| \\
&\quad + \left|\int_{[0,1]^m}\sum_{i=1}^m g_i\big(F_{\mu_i}(x_i)\big)\,\big(d\nu^n(x)-d\nu(x)\big)\right| \\
&\le \int_{[0,1]^m}\sum_{i=1}^m \left|g_i(F_{\mu_i^n}(x_i))-g_i(F_{\mu_i}(x_i))\right|\,d\nu^n(x) \\
&\quad + \left|\int_{[0,1]^m}\sum_{i=1}^m g_i\big(F_{\mu_i}(x_i)\big)\,\big(d\nu^n(x)-d\nu(x)\big)\right| \\
&\le \int_{[0,1]^m}\sum_{i=1}^m q\left|F_{\mu_i^n}(x_i)-F_{\mu_i}(x_i)\right|\,d\nu^n(x) \\
&\quad + \left|\int_{[0,1]^m}\sum_{i=1}^m g_i\big(F_{\mu_i}(x_i)\big)\,\big(d\nu^n(x)-d\nu(x)\big)\right| \\
&\le mq\max_{1\le i\le m}\left\|F_{\mu_i^n}-F_{\mu_i}\right\|_\infty \\
&\quad + \left|\int_{[0,1]^m}\sum_{i=1}^m g_i\big(F_{\mu_i}(x_i)\big)\,\big(d\nu^n(x)-d\nu(x)\big)\right|.
\end{aligned}
$$

Since $\mu_i^n \to \mu_i$ in distribution and μ_i is continuous, then the convergence of F_{μ_i} to F_μ is uniform. The first term vanishes as $n \to \infty$, and since $\nu_i^n \to \nu_i$ in distribution, the second term also vanishes. Which proves the joint continuity of g. □

One can now apply Berge's maximum theorem with $\mathcal{X} = \mathcal{Y} = E_m^c$. In the present setting, F is the constant correspondence equal to E_m^c ; the assumptions are then immediately satisfied, and the theorem guarantees that, for every $\mu \in E_m^c$, one has $\Psi(\mu) \neq \varnothing$ and that the graph of Ψ is closed. Moreover, since the objective is linear in ν, the set $\Psi(\mu)$ is convex.
This in turn makes it possible to apply Kakutani's fixed point theorem and to solve this simplified version of the mean field problem.

Bibliographie

[1] Ahmadinejad, A., Dehghani, S., Hajiaghayi, M., Lucier, B., Mahini, H., Seddighin, S. : From duels to battlefields : Computing equilibria of blotto and other games. Mathematics of Operations Research **44**(4), 1304–1325 (2019). DOI 10.1287/moor.2018.0971. URL https://doi.org/10.1287/moor.2018.0971

[2] Ali Khan, M., Rustichini, A. : Cournot-nash equilibrium distributions of games with uncertainty and imperfect information. Journal of Mathematical Economics **22**(1), 35–59 (1993). DOI https://doi.org/10.1016/0304-4068(93)90029-K. URL https://www.sciencedirect.com/science/article/pii/030440689390029K

[3] Aliprantis, C.D., Border, K.C. : Infinite Dimensional Analysis : A Hitchhiker's Guide, 3rd edn. Springer-Verlag Berlin Heidelberg. Springer (2006). DOI 10.1007/3-540-29587-9

[4] Beaglehole, D., Hopkins, M., Kane, D., Liu, S., Lovett, S. : Sampling equilibria : Fast no-regret learning in structured games. In : Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), SODA'23, pp. 3817–3855. Society for Industrial and Applied Mathematics, Philadelphia, PA (2023). DOI 10.1137/1.9781611977554.ch149. URL https://doi.org/10.1137/1.9781611977554.ch149

[5] Beale, E.M.L., Heselden, G.P.M. : An approximate method of solving blotto games. Naval Research Logistics Quarterly **9**(2), 65–79 (1962). DOI 10.1002/nav.3800090202. URL https://ideas.repec.org/a/wly/navlog/v9y1962i2p65-79.html

[6] Behnezhad, S., Dehghani, S., Derakhshan, M., Hajiaghayi, M., Seddighin, S. : Fast and simple solutions of blotto games. Operations Research (2022). DOI 10.1287/opre.2022.2261. URL https://doi.org/10.1287/opre.2022.2261. Articles in Advance

[7] Benamou, J.D., Carlier, G., Cuturi, M., Nenna, L., Peyré, G. : Iterative bregman projections for regularized transportation problems. SIAM Journal on Scientific Computing **37**(2), A1111–A1138 (2015). DOI 10.1137/141000439. URL https://doi.org/10.1137/141000439

[8] Berge, C. : Topological Spaces : Including a Treatment of Multi-Valued Functions, Vector Spaces and Convexity. Oliver and Boyd, Edinburgh (1963)

[9] Blanchet, A., Carlier, G. : From nash to cournot–nash equilibria via the monge–kantorovich problem. Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences **372**(2028), 20130398 (2014). DOI 10.1098/rsta.2013.0398. URL https://doi.org/10.1098/rsta.2013.0398

[10] Blanchet, A., Carlier, G. : Optimal transport and cournot-nash equilibria. Math. Oper. Res. **41**(1), 125–145 (2016). DOI 10.1287/moor.2015.0719. URL https://doi.org/10.1287/moor.2015.0719

[11] Boix-Adserà, E., Edelman, B., Jayanti, S. : The multiplayer Colonel Blotto game. Games and Economic Behavior **129**, 15–31 (2021)

[12] Émile Borel : La théorie du jeu et les équations intégrales à noyau symétrique. Comptes Rendus de l'Académie des Sciences **173**, 1304–1308 (1921). URL https://gallica.bnf.fr/ark:/12148/bpt6k31267/f1304.item.r=Comptes+rendus+hebdomadaires+des+s%C3%A9ances+de+l.langFR

[13] Émile Borel, Ville, J. : Application de la théorie des probabilités aux jeux de hasard : cours professé à la Faculté des sciences de Paris, *Traité du calcul des probabilités et de ses applications*, vol. 4. Gauthier-Villars, Paris (1938)

[14] Briand, P., Hu, Y. : Bsde with quadratic growth and unbounded terminal value. Probability Theory and Related Fields **136**(4), 604–618 (2006). DOI 10.1007/s00440-006-0497-0. URL https://doi.org/10.1007/s00440-006-0497-0

[15] Caines, P.E., Huang, M., Malhamé, R.P. : Mean Field Games, pp. 1–28. Springer International Publishing, Cham (2017). DOI 10.1007/978-3-319-27335-8_7-1. URL https://doi.org/10.1007/978-3-319-27335-8_7-1

[16] Carmona, R. : Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications. SIAM Financial Mathematics Series. SIAM (2016). DOI 10.1137/1.9781611974249

[17] Carmona, R., Delarue, F. : Probabilistic Theory of Mean Field Games with Applications I-II, *Probability Theory and Stochastic Modelling*, vol. 83-84. Springer (2018). DOI 10.1007/978-3-319-58920-6

[18] Chen, F., Ren, Z., Wang, S. : Entropic Fictitious Play for Mean Field Optimization Problem. Journal of Machine Learning Research **24**, no. 211 (2023). URL https://ip-paris.hal.science/hal-04508395

[19] Crandall, M.G., Ishii, H., Lions, P.L. : user's guide to viscosity solutions of second order partial differential equations (1992). URL https://arxiv.org/abs/math/9207212

[20] Cuturi, M. : Sinkhorn distances : Lightspeed computation of optimal transport. In : Advances in Neural Information Processing Systems (2013)

[21] Cvitanić, J., Possamaï, D., Touzi, N. : Moral hazard in dynamic risk management. Management Science **63**(10), 3328–3346 (2017). DOI 10.1287/mnsc.2016.2493

[22] Cvitanić, J., Possamaï, D., Touzi, N. : Dynamic programming approach to principal-agent problems. Finance and Stochastics **22**(1), 1–37 (2018)

[23] De Silva, L. : On the efficiency of decentralized epidemic management and competitive viral marketing. Ph.D. thesis, Université de Lorraine (2023). URL http://docnum.univ-lorraine.fr/public/DDOC_T_2023_0145_LINDAMULAGE_DE_SILVA.pdf

[24] Dziubiński, M. : The spectrum of equilibria for the Colonel Blotto and the Colonel Lotto games. In : V. Bilò, M. Flammini (eds.) Algorithmic Game Theory, *Lecture Notes in Computer Science*, vol. 10504, pp. 292–306. Springer, Cham (2017)

[25] Dziubiński, M. : Non-symmetric discrete colonel blotto game (2025). URL https://arxiv.org/abs/2511.10827

[26] E, W., Han, J., Jentzen, A. : Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. Communications in Mathematics and Statistics **5**(4), 349–380 (2017). DOI 10.1007/s40304-017-0117-6. URL http://dx.doi.org/10.1007/s40304-017-0117-6

[27] Elie, R., Possamaï, D. : Contracting theory with competitive interacting agents. SIAM Journal on Control and Optimization **57**(2), 1157–1188 (2019)

[28] Ferdowsi, A., Saad, W., Maham, B., Mandayam, N.B. : A colonel blotto game for interdependence-aware cyber-physical systems security in smart cities. In : Proceedings of the 2nd International Workshop on Science of Smart City Operations and Platforms Engineering (SCOPE 2017), pp. 7–12. ACM, Pittsburgh, PA, USA (2017). DOI 10.1145/3063386.3063765

[29] Ferdowsi, A., Saad, W., Mandayam, N.B. : Colonel blotto game for secure state estimation in interdependent critical infrastructure. CoRR **abs/1709.09768** (2017). URL `https://arxiv.org/abs/1709.09768`

[30] Fiedler, M. : Algebraic connectivity of graphs. Czechoslovak Mathematical Journal **23**(2), 298–305 (1973). URL `https://dml.cz/bitstream/handle/10338.dmlcz/101168/CzechMathJ_23-1973-2_11.pdf`

[31] Gross, O., Wagner, R. : A continuous colonel blotto game. RAND Corporation Research Memorandum **RM-408** (1950)

[32] Grossman, S.J., Hart, O.D. : An analysis of the principal-agent problem. Econometrica **51**(1), 7–45 (1983)

[33] Guan, S., Wang, J., Yao, H., Jiang, C., Han, Z., Ren, Y. : Colonel blotto games in network systems : Models, strategies, and applications. IEEE Transactions on Network Science and Engineering **7**(2), 637–649 (2019)

[34] Gupta, A., Schwartz, G., Langbort, C., Sastry, S.S., Bařar, T. : A three-stage colonel blotto game with applications to cyberphysical security. In : 2014 American Control Conference, pp. 3820–3825. IEEE (2014)

[35] Hajimirsaadeghi, M., Mandayam, N.B. : A dynamic colonel blotto game model for spectrum sharing in wireless networks. In : 2017 55th Annual Allerton conference on communication, control, and computing (Allerton), pp. 287–294. IEEE (2017)

[36] Hamadene, S., Lepeltier, J. : Backward equations, stochastic control and zero-sum stochastic differential games. Stochastics and Stochastic Reports **54**(3-4), 221–231 (1995). DOI 10.1080/17442509508834006

[37] Han, J., Jentzen, A., E, W. : A brief review of the deep bsde method for solving high-dimensional partial differential equations (2025). DOI 10.48550/arXiv.2505.17032. URL `https://arxiv.org/abs/2505.17032`. ICBS proceedings of Frontiers of Science Awards (2024)

[38] Hart, S. : Discrete Colonel Blotto and General Lotto games. International Journal of Game Theory **36**, 441–460 (2008)

[39] Hazzami, A., Kazi-Tani, N., Varma, V.S. : Zero-Sum Stochastic Differential Blotto-Type Games on Graphs (2025). URL `https://hal.univ-lorraine.fr/hal-05416927`. Working paper or preprint

[40] Hazzami, A., Kazi-Tani, N., Varma, V.S. : Modified deep-bsde solver for a stochastic differential blotto game (2026). DOI 10.5281/zenodo.18303818. URL `https://github.com/MohamedAmineHazzami/Modified-Deep-BSDE`. GitHub repository ; archived on Zenodo.

[41] Hernández, D.G., Zanette, D.H. : Evolutionary dynamics of resource allocation in the colonel blotto game. Journal of Statistical Physics **151**(3–4), 623–636 (2013). DOI 10.1007/s10955-012-0659-7

[42] Hernández-Santibáñez, N., Mastrolia, T. : Contract theory in a vuca world. SIAM Journal on Control and Optimization **57**(4), 3072–3100 (2019)

[43] Hernández-Santibánez, N., Mastrolia, T. : Contract theory in a vuca world. SIAM Journal on Control and Optimization **57**(4), 3072–3100 (2019). DOI 10.1137/18M1184527. URL `https://doi.org/10.1137/18M1184527`

[44] Hernández-Santibáñez, N. : Principal-multiagents problem under equivalent changes of measure : General study and an existence result. Stochastic Processes and their Applications **177**, 104448 (2024). DOI https://doi.org/10.1016/j.spa.2024.104448. URL `https://www.sciencedirect.com/science/article/pii/S0304414924001546`

[45] Holmström, B. : Moral hazard and observability. The Bell Journal of Economics **10**(1), 74–91 (1979)

[46] Huang, M., Malhamé, R., Caines, P.E. : Large population stochastic dynamic games : closed-loop mckean–vlasov systems and the nash certainty equivalence principle. Communications in Information & Systems **6**(3), 221–252 (2006)

[47] Jevons, W.S. : The Theory of Political Economy. Macmillan and Co. (1871)

[48] Kakutani, S. : A generalization of brouwer's fixed point theorem. Duke Mathematical Journal **8**(3), 457–459 (1941). DOI 10.1215/S0012-7094-41-00838-4

[49] Karoui, N.E., Peng, S., Quenez, M.C. : Backward stochastic differential equations in finance. Mathematical Finance **7**(1), 1–71 (1997). URL `https://EconPapers.repec.org/RePEc:bla:mathfi:v:7:y:1997:i:1:p:1-71`

[50] Kovenock, D., Roberson, B. : Conflicts with multiple battlefields. Mathematics of Operations Research **37**(1), 1–20 (2011). DOI 10.1287/moor.1110.0536

[51] Kovenock, D., Roberson, B. : Coalitional colonel blotto games with application to the economics of alliances. Journal of Public Economic Theory **14**(4), 653–676 (2012). DOI 10.1111/j.1467-9779.2012.01556.x

[52] Kovenock, D., Roberson, B. : Generalizations of the general lotto and colonel blotto games. Economic Theory **71**(3), 997–1032 (2021)

[53] Lasry, J.M., Lions, P.L. : Jeux à champ moyen. i–le cas stationnaire. Comptes Rendus Mathématique **343**(9), 619–625 (2006). DOI 10.1016/j.crma.2006.09.019

[54] Lasry, J.M., Lions, P.L. : Mean field games. Japanese Journal of Mathematics **2**(1), 229–260 (2007). DOI 10.1007/s11537-007-0657-8

[55] Leininger, W. : A generalization of the 'maximum theorem'. Economics Letters **15**(3–4), 309–313 (1984)

[56] Liang, D., Wang, Y., Cao, Z., Yang, X. : Discrete Colonel Blotto games with two battlefields. International Journal of Game Theory **52**(4), 1111–1151 (2023)

[57] Ma, J., Protter, P., Yong, J. : Solving forward-backward stochastic differential equations explicitly—a four step scheme. Probability Theory and Related Fields **98**(3) (1994)

[58] Macdonell, S., Mastronardi, N. : Waging simple wars : A complete characterization of two-battlefield blotto equilibria. Economic Theory **58**(1) (2015). URL `https://ssrn.com/abstract=2762255`. Available at SSRN : 2762255

[59] Mastrolia, T., Possamaï, D. : Moral hazard under ambiguity. Journal of Optimization Theory and Applications **179**(2), 452–500 (2018)

[60] Nash, J.F. : Equilibrium points in n-person games. Proceedings of the National Academy of Sciences **36**(1), 48–49 (1950). DOI 10.1073/pnas.36.1.48

[61] von Neumann, J., Morgenstern, O. : Theory of Games and Economic Behavior. Princeton University Press (1944)

[62] Owen-Smith, N. : Allocation of food gains to storage versus growth in a variable environment. International Congress Series **1275**, 321–326 (2004). DOI https://doi.org/10.1016/j.ics.2004.08.079. URL https://www.sciencedirect.com/science/article/pii/S0531513104014645. Animals and Environments. Proceedings of the Third International Conference of Comparative Physiology and Biochemistry

[63] Pardoux, E., Peng, S. : Adapted solution of a backward stochastic differential equation. Systems & Control Letters **14**(1), 55–61 (1990). DOI https://doi.org/10.1016/0167-6911(90)90082-6. URL https://www.sciencedirect.com/science/article/pii/0167691190900826

[64] Pareto, V. : Cours d'économie politique. F. Rouge (1896)

[65] Peng, S. : Stochastic hamilton–jacobi–bellman equations. SIAM J. Control Optim. **30**(2), 284–304 (1992)

[66] Penn, A.I. : A generalized lagrange-multiplier method for constrained matrix games. Operations Research **19**(4), 933–945 (1971). URL http://www.jstor.org/stable/169052

[67] Raissi, M. : Forward-backward stochastic neural networks : Deep learning of high-dimensional partial differential equations (2018). URL https://arxiv.org/abs/1804.07010

[68] Roberson, B. : The colonel blotto game. Economic Theory **29**(1), 1–24 (2006). URL https://EconPapers.repec.org/RePEc:spr:joecth:v:29:y:2006:i:1:p:1-24

[69] Roberson, B., Kvasov, D. : The non-constant-sum Colonel Blotto game. Economic Theory **51**(2), 397–433 (2012)

[70] Sannikov, Y. : A continuous-time version of the principal-agent problem. The Review of Economic Studies **75**(3), 957–984 (2008)

[71] Schmeidler, D. : Equilibrium points of nonatomic games. Journal of Statistical Physics **7**(4), 295–300 (1973). DOI 10.1007/BF01014905. URL https://doi.org/10.1007/BF01014905

[72] Schwalbe, U., Walker, P. : Zermelo and the early history of game theory. Mathematical Social Sciences **34**(2), 111–138 (1997). DOI 10.1016/S0165-4896(97)00003-X. Includes an English translation of Zermelo's 1913 paper

[73] Shavell, S. : Risk sharing and incentives in the principal and agent relationship. The Bell Journal of Economics **10**(1), 55–73 (1979)

[74] Smith, A. : An Inquiry into the Nature and Causes of the Wealth of Nations. W. Strahan and T. Cadell, London (1776)

[75] Spielman, D.A. : Spectral and algebraic graph theory. http://cs-www.cs.yale.edu/homes/spielman/sagt (2025). Incomplete draft, dated April 2, 2025

[76] Sung, J. : Contract Theory : Discrete- and Continuous-Time Models. Springer, Singapore (2023). DOI 10.1007/978-981-99-5487-2

[77] Touzi, N. : Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE, vol. 29 (2013). DOI 10.1007/978-1-4614-4286-8

[78] Tzu, S. : L'art de la guerre. Larousse (2008). Traduction française, édition commentée

[79] Vu, D.Q. : Colonel blotto game – resource allocation. https://dongquan-vu.github.io/interests/ (2023). Accessed August 29, 2025

[80] Vu, D.Q., Loiseau, P. : Colonel blotto games with favoritism : Competitions with pre-allocations and asymmetric effectiveness. In : Proceedings of the 22nd ACM Conference on

Economics and Computation, EC'21, pp. 862–863. Association for Computing Machinery, New York, NY, USA (2021). DOI 10.1145/3465456.3467527. URL `https://doi.org/10.1145/3465456.3467527`

[81] Walras, L. : Elements of pure economics : Or the theory of social wealth. Elements of Pure Economics : Or the Theory of Social Wealth pp. 1–620 (2013). DOI 10.4324/9781315888958

[82] Zermelo, E. : Über eine anwendung der mengenlehre auf die theorie des schachspiels. Proceedings of the Fifth International Congress of Mathematicians **2**, 501–504 (1913)

Abstract

This thesis studies resource allocation games in general, and the Colonel Blotto game in particular, in three settings : stochastic differential games, continuous-time principal–agent contracting, and mean-field (infinite-population) interactions.

Chapter 2 introduces a stochastic differential generalization of the classical Blotto game. Two players allocate budgets across multiple battlefields while the state evolves according to a controlled stochastic differential system, allowing for non-trivial cross-battlefield interdependencies, for instance through spectral objectives in the players' payoff functionals. Using a dynamic programming approach, we provide an explicit characterization of pure-strategy Nash equilibria, which take a bang–bang form. In the symmetric case, equilibria are characterized through a parabolic PDE ; in the asymmetric case, they are characterized through a backward stochastic differential equation (BSDE). We also propose a numerical scheme to solve the resulting BSDEs and illustrate the equilibrium behavior.

Chapter 3 extends the Blotto-type game to a contract-theoretic principal–agent framework. Adopting a Stackelberg viewpoint, a principal designs both a budget process and a terminal contract for an agent who then competes against a hacker in a stochastic differential Blotto-type game. Since neither the principal nor the agent observes the hacker's objective, we adopt a worst-case viewpoint. A key contribution is a *collapse* result that removes one layer of contract design : although the hacker is contracted by an external entity, we show that, under a rich class of admissible contracts, the principal's worst-case evaluation over hacker contracts is equivalent to an adverse choice of the hacker's action. This reduction yields a tractable Isaacs (HJBI) characterization of the principal's worst-case value. We further characterize pure-strategy Nash equilibria of the agent–hacker game, again of bang–bang type, and provide numerical illustrations in the same spirit as Chapter 2.

Finally, Chapter 4 studies an infinite-population extension of the Blotto game. The resulting formulation becomes a fixed-point problem on a space of probability measures over the simplex. In its general form, the model admits both trivial equilibria in the form of Dirac masses and equilibria with densities (absolutely continuous with respect to Lebesgue measure). To introduce more structure, we incorporate an entropic penalization and define a mean-field version applicable to a broad class of resource allocation problems. In this setting, an agent's best response is characterized explicitly as a Gibbs measure, and a fixed-point theorem yields existence of a mean-field equilibrium. For Blotto-type rewards, we further provide an explicit construction reducing the measure-valued fixed-point problem to a finite-dimensional one.

Keywords: Resource allocation games, stochastic differential games, backward stochastic differential equations (BSDE), principal–agent problems, infinite-population games.

Résumé

Cette thèse étudie les jeux d'allocation de ressources en général, et le jeu du Colonel Blotto en particulier, dans trois cadres : les jeux différentiels stochastiques, la théorie des contrats en temps continu (principal–agent), et les interactions de type champ moyen (population infinie).

Le Chapitre 2 introduit une généralisation du jeu de Blotto en cadre stochastique différentiel. Deux joueurs répartissent leurs budgets sur plusieurs champs de bataille, tandis que le processus d'état évolue selon un système différentiel stochastique contrôlé, permettant des interdépendances non triviales entre champs de bataille, notamment lorsque les objectifs des joueurs font intervenir des critères spectraux. Par une approche de programmation dynamique, on donne une caractérisation explicite d'équilibres de Nash en stratégies pures, qui sont de type bang–bang. Dans le cas symétrique, les équilibres sont caractérisés via une EDP parabolique ; dans le cas asymétrique, ils sont caractérisés via une équation différentielle stochastique rétrograde (EDSR). On propose également un schéma numérique pour résoudre les EDSR obtenues et illustrer le comportement à l'équilibre.

Le Chapitre 3 étend un jeu d'allocation de ressources de type Blotto à un cadre de théorie des contrats principal–agent. Dans une perspective de Stackelberg, un principal conçoit à la fois un processus de budget et un contrat terminal pour un agent. Ce dernier affronte ensuite un hacker dans un jeu de type Blotto formulé comme jeu différentiel stochastique. Comme ni le principal ni l'agent n'observent l'objectif du hacker, on adopte un point de vue au pire cas. Une contribution clé est un résultat de collapse (élimination) qui supprime un niveau de conception contractuelle : bien que le hacker soit engagé par une entité externe, on montre que, sous une classe riche de contrats admissibles, l'évaluation au pire cas du principal sur les contrats du hacker est équivalente à un choix adverse direct de l'action du hacker. Cette réduction conduit à une caractérisation de la valeur au pire cas du principal par une équation d'Isaacs (HJBI) et par une équation différentielle stochastique rétrograde EDSR. On caractérise en outre des équilibres de Nash en stratégies pures pour le jeu agent–hacker, là encore de type bang–bang, et on propose des illustrations numériques dans le même esprit que le Chapitre 2.

Enfin, le Chapitre 4 étudie une extension du jeu de Blotto à une population infinie. La formulation obtenue se ramène à un problème de point fixe sur un espace de mesures de probabilité sur le simplexe. Dans sa forme générale, le modèle admet à la fois des équilibres triviaux sous forme de masses de Dirac et des équilibres possédant une densité (absolument continue par rapport à la mesure de Lebesgue). Afin d'introduire davantage de structure, on utilise une pénalisation entropique et on définit une version champ moyen applicable à une large classe de problèmes d'allocation de ressources. Dans ce cadre, la meilleure réponse d'un agent face à la population est caractérisée explicitement comme une mesure de Gibbs, et un théorème de point fixe assure l'existence d'un équilibre champ moyen. Dans le cas de gains de type Blotto, on fournit de surcroît une construction explicite qui réduit le problème de point fixe sur un espace de mesures à un problème en dimension finie.

Mots-clés: Jeux d'allocation de ressources, jeux différentiels stochastiques, équations différentielles stochastiques rétrogrades (EDSR), problèmes principal–agent, jeux à population infinie.