PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à l'Université Paris-Dauphine

Systemic Risk, Complex Financial Networks and Graphon Mean Field Interacting Systems

Soutenue par

Zhongyuan CAO

Le 22 Septembre 2023

École doctorale n°543

Ecole doctorale SDOSE

Spécialité

Mathématiques appliquées

Composition du jury :

M. Aurélien ALFONSI Professeur, Ecole des Ponts Paristech	*Examinateur*
M. Hamed AMINI Associate professor, University of Florida	*Directeur de thèse*
M. Erhan BAYRAKTAR Professor, University of Michigan	*Rapporteur*
M. Rama CONT Professor, University of Oxford	*Rapporteur*
M. Laurent MASSOULIE Directeur de recherche, INRIA	*Président du jury*
M. Huyên PHAM Professeur des universités, Université Paris Cité	*Examinateur*
Mme Agnès SULEM Directeur de recherche, INRIA	*Directrice de thèse*
Mme Luitgard VERAART Professor, The London School of Economics and Political Science	*Examinatrice*

Dauphine UNIVERSITÉ PARIS | PSL

Acknowledgements

First of all, I'd like to give my most sincere and deepest gratitude to my advisors, Prof. Hamed Amini and Prof. Agnès Sulem. During my second year of master's studies, Prof. Agnès Sulem provided a research intern at INRIA to me. This was the first time I got close to mathematical research. She proposed a very interesting topic to me, which was "systemic risk and mean field games" and later extended to be my research topics in my Ph.D. study. After my research intern, I was so lucky to have the opportunity to start a thesis under the supervision by her and Prof. Hamed Amini. They are so kind and both are very good supervisors. Hamed always proposed innovative and interesting ideas, discussed with me in very detailed stuff. They were passionate during our discussions and never stopped giving me advice to help me improve and polish papers. When I was faced with any obstacles, they were always there. Throughout the past three years, their inspiring insights and helpful advice were invaluable to me. Their strict and ceaseless supervision on my works have been motivating me until the completion of this thesis. It's a so special and unique experience to learn so much from them.

During the three years, I am profoundly grateful to my host institution, INRIA, especially the research team Mathrisk, where I did my Ph.D. study and completed this thesis. Mathrisk is a warm big family. I got lots of valuable things and met many kind and interesting people here. INRIA provideds me with financial supporting for my academic travels. It allowed me to have the opportunity to participate in different conferences, keeping up with the newest research trends and communicating with other researchers. I am so lucky to have the financial supporting from DIM Math Innov. I am so grateful for that. Without this, I could not even start the thesis. I also want to thank my host institution. I also want to express my thanks to Université Paris-Dauphine, to CEREMADE for providing incredible environment for research.

It is an honour to have Prof. Erhan Bayraktar and Prof. Rama Cont as my thesis reviewers. I extend my sincere thanks to them for reviewing my thesis carefully and giving me many inspirational and helpful suggestions and comments to the thesis. These comments helped me improve my thesis and motivated me to think more and deeper in my research and future works. I' d like to express my deep respect and profound gratitude to all members of my thesis committee, Prof. Aurélien Alfonsi, Prof. Erhan Bayraktar, Prof. Rama Cont, Prof. Laurent Massoulié, Prof. Huyên Pham and Prof. Luitgard Veraart. It is an honour to have you in my thesis committee. I am very glad to be able to hear your comments to my thesis. I'd like to give special thanks to Prof. Aurélien Alfonsi for his kind and helpful discussion upon the project *Fire sales and default cascade*, in Chapter 3 of this thesis. I am also grateful for having the opportunity to collaborate with Prof. Andreea Minca in the project *ruin probabilities for risk processes in stochastic networks*, in Chapter 4. I'd like to express my sincere gratitude to Prof. Mathieu Laurière. We discussed a lot on mean field games and I learned a lot from him.

Taking this moment, I would like to express my deepest gratitude to my girlfriend Wanqing. We were schoolmates in bachelor study. Since then, we started to communicate and discuss a lot in mathematics, and also encouraged each other to go further in the path of math study. We have the same dream and keep going in the same direction. Most importantly, it is her who encouraged me to come to France, to have the opportunity to accept the very top math practice in the world, establishing a solid math base. She impelled me to have the nerve, the tenacity, and the confidence to explore the world, and the world of mathematics.

I'd like to give endless thanks to my family, especially my parents in China, for everything that I can imagine. They have supported me with love and understanding. They provided me with unconditional support both materially and spiritually. In particular, for my first two years in France, when I was in master study. I had no financial supporting at that moment. My parents provided a huge amount of money to cover all my cost to live and study in such a high living cost city, Paris. Every time when I encountered the obstacle that made me feel dejected, they were always there listening to me, regardless of the time zone difference.

I am very grateful to have so many warm-hearted and reliable friends and colleagues, Wenye, Zhuoying, Z. Alexandre, Jiamin, Haibo, Kexin, Jiayi, Qi...., accompanying me during these years. Special thanks to Haibo for taking care of my cats when I am on vacation. Endless and sincere thanks to Wenye and Zhuoying for tasty dinners so many times in your cozy and sweet house. Every time when I could not continue my thesis, I miss your dinner, just like I miss the dinner in my grandmother's house!

At last, I'd like to mention my cute cat named "profiterole" in French, also with Chinese name "Fufu". She is a very considerate little cat who has given me a lot of companionship and joy. Sincere thanks to her!

Coming along all the way, to every single individual who has ever shown your kindness, to every little heartwarming act in my life, to every bit of landscape scoring the beauty of world. Please allow me to give my most profound thanks!

The journey is not over, this is just a new beginning, the afterglow, the starry sky, the dawn, the bird's song, moving, keep moving, moving forward.....

Contents

Résumé Détaillé **i**

1 Introduction **1**

1.1 Financial networks and random graphs . . . 1

1.1.1 Graphs, degrees and connectivity structure . . . 2

1.1.2 Random graph with given vertex degrees . . . 3

1.2 Default cascade and fire sales . . . 4

1.2.1 Threshold model and bootstrap percolation . . . 4

1.2.2 Default cascade and asymptotic results . . . 5

1.2.3 Fire sales and equilibrium price . . . 9

1.3 Risk processes on networks . . . 12

1.3.1 Classical risk processes . . . 13

1.3.2 Risk processes on financial networks . . . 13

1.4 Graphon mean field games and interacting systems . . . 17

1.4.1 Graphon mean field backward stochastic differential equations with jumps . . . 19

1.4.2 Stochastic games with graphon mean field interactions . . . 26

1.4.3 Graphon equilibria and approximate Nash equilibria . . . 27

1.5 Publications and working papers . . . 32

I SYSTEMIC RISK MODELS IN COMPLEX FINANCIAL NETWORKS **34**

2 Limit Theorems for Default Contagion and Systemic Risk **35**

2.1 Introduction . . . 36

2.2 Model . . . 39

2.2.1 Financial Network and Default Cascade . . . 39
2.2.2 Node Classification and Configuration Model . . . 40
2.2.3 Death Process and Final Solvent Institutions . . . 41
2.3 Limit Theorems . . . 43
2.3.1 Asymptotic Magnitude of Default Contagion . . . 44
2.3.2 Asymptotic Normality of Default Contagion . . . 46
2.4 Quantifying Systemic Risk . . . 49
2.4.1 Numerical analysis on heterogeneity and stability . . . 51
2.5 Targeting Interventions in Financial Networks . . . 52
2.6 Default cascade process . . . 54
2.7 Proof of Main Theorems . . . 56
2.7.1 Proof of Theorem 2.1 . . . 56
2.7.2 Proof of Theorem 2.4 . . . 57
2.7.3 Proof of Theorem 2.6 . . . 58
2.7.4 Proof of Theorem 2.7 . . . 59
2.7.5 Proof of Theorem 2.8 . . . 62
2.7.6 Proof of Theorem 2.9 . . . 63
2.7.7 Proof of Theorem 2.10 . . . 64
2.7.8 Proofs in §2.6 . . . 65
2.8 Appendix . . . 70
2.8.1 Proof of Lemmas . . . 70
2.8.2 Proof of Lemma 2.2 . . . 71
2.8.3 Proof of Lemma 2.13 . . . 71
2.8.4 Proof of Lemma 2.14 . . . 73
2.8.5 Proof of Lemma 2.15 . . . 78
2.8.6 Proof of Lemma 2.16 . . . 79
2.8.7 Covariances in Theorem 2.7 . . . 83
2.8.8 Covariances in Theorem 2.9 . . . 86
2.9 Concluding Remarks . . . 87

3 Fire Sales and Default Cascades **89**

3.1 Introduction . . . 91
3.2 The Model . . . 94
3.2.1 Financial network . . . 94
3.2.2 Default cascade . . . 96
3.2.3 Node classification . . . 97
3.2.4 A simple model of fire sales . . . 99
3.3 Limit Theorems . . . 101
3.3.1 Asymptotic magnitude of default cascade with fire sales . . . 102
3.3.2 Asymptotic normality of default cascade with fire sales . . . 105
3.4 Numerical Experiments . . . 109
3.4.1 Numerical set-up . . . 110
3.4.2 Regular financial networks . . . 111
3.4.3 Core-Periphery financial networks . . . 113
3.4.4 Scale-Free financial networks . . . 115
3.4.5 Asymptotic normality and confidence intervals . . . 117
3.5 Proofs . . . 121
3.5.1 Auxiliary lemmas . . . 121
3.5.2 Proof of Theorem 3.7 . . . 124
3.5.3 Proof of Theorem 3.9 . . . 127
3.5.4 Proof of Theorem 3.11 . . . 127
3.5.5 Discussion on Lemma 3.6 and Lemma 3.13 . . . 128
3.5.6 Proof of Theorem 3.14 . . . 130
3.5.7 Proof of Theorem 3.16 . . . 136
3.5.8 Proof of Theorem 3.18 . . . 138
3.5.9 Proof of Theorem 3.19 . . . 139
3.6 Extension to Multiple Illiquid Assets . . . 142
3.7 Concluding Remarks . . . 149

4 Ruin probabilities for Risk Processes in Stochastic Networks **151**
4.1 Introduction . . . 152
4.2 The Model . . . 154

4.2.1 Networked risk processes . . . 155
4.2.2 Node classification . . . 155
4.2.3 The loss reveal process . . . 157
4.3 Main Theorems . . . 158
4.3.1 Asymptotic analysis of the general loss reveal process . . . 158
4.3.2 Ruin probabilities for the networked risk processes . . . 163
4.4 Proofs . . . 165
4.4.1 Some preliminary results . . . 165
4.4.2 Proof of Lemma 4.2 . . . 166
4.4.3 Proof of Theorem 4.5 . . . 169
4.4.4 Proof of Theorem 4.7 . . . 172
4.4.5 Proof of Lemma 4.8 . . . 173
4.4.6 Proof of Theorem 4.9 . . . 173
4.4.7 Proof of Theorem 4.10 . . . 173
4.4.8 Proof of Lemma 4.11 . . . 177
4.5 Complex Networked Risk Processes . . . 177
4.6 Concluding Remarks . . . 180

II GRAPHON MEAN FIELD INTERACTING SYSTEMS 182

5 Graphon Mean Field Backward Stochastic Differential Equations and Associated Dynamic Risk Measures 183

5.1 Introduction . . . 184
5.2 Graphon mean-field BSDEs with jumps . . . 185
5.2.1 Notation and definitions . . . 185
5.2.2 Graphons . . . 186
5.2.3 Existence and uniqueness results . . . 188
5.2.4 Comparison theorems . . . 195
5.2.5 Continuity and stability results . . . 197
5.3 Convergence of interacting particle systems to graphon mean-field BSDEs . . . 204
5.4 Graphon dynamic risk measures . . . 212
5.4.1 Definition and properties . . . 212

5.4.2 Dual representation 214
5.5 Concluding remarks 223

6 Stochastic Graphon Mean Field Games with Jumps and Approximate Nash Equilibria **225**

6.1 Introduction 227
6.2 Probabilistic set-up and notations 228
6.3 Graphon mean field games with jumps 230
6.4 Finite networked games with heterogenous interactions 234
6.4.1 Finite game with jumps 235
6.4.2 Propagation of chaos for controlled graphon system 236
6.5 Approximate Nash equilibria of finite games 239
6.6 Proofs 242
6.6.1 Existence of graphon equilibrium: Proof of Theorem 6.4 242
6.6.2 Stability of graphon: Proof of Theorem 6.8 248
6.6.3 Proof of Lemma 6.9 250
6.6.4 Large population convergence: Proof of Theorem 6.10 251
6.6.5 Convergence for general graphon: Proof of Corollary 6.13 255
6.6.6 Stability of control: Proof of Lemma 6.17 256
6.6.7 Proof of Theorem 6.18 and Theorem 6.20 259
6.7 Concluding remarks 263

Résumé Détaillé

Cette thèse est divisée en deux parties. La première partie étudie la stabilité et le risque systémique des grands réseaux financiers complexes, notamment la contagion de défaut, le processus de risque en réseau et un modèle de ventes forcées. Nous commençons par prouver certains théorèmes limite (lois des grands nombres et théorèmes de la limite centrale) concernant la dynamique de contagion. Nous montrons comment quantifier le risque systémique pour un réseau financier en présence d'informations partielles face à un choc extérieur. Ensuite, nous présentons un cadre général et accessible pour comprendre l'impact conjoint des ventes forcées et des cascades de défauts sur le risque systémique dans les réseaux financiers complexes. Enfin, nous étudions les processus de risque pour les grands systèmes financiers, où les agents, situés sur un vaste réseau, subissent des pertes de la part de leurs voisins.

La deuxième partie de la thèse se concentre sur les systèmes d'interaction de champ moyen avec sauts et les jeux de champ moyen avec graphon. Ici, le réseau financier est considéré comme un vaste système d'interaction, avec une structure de champ moyen dépendant de la structure du graphe sous-jacent du réseau. Nous commençons par mener une étude approfondie des équations différentielles stochastiques rétrogrades de champ moyen avec sauts (EDSR) et des mesures de risque dynamiques globales associées. Ensuite, nous étudions les jeux stochastiques continus avec des interactions de champ moyen inhomogènes sur de vastes réseaux et examinons leurs limites graphon. Nous proposons des équilibres de Nash approximatifs pour les jeux finis avec des interactions hétérogènes, en utilisant leurs équilibres graphon comme références.

Réseaux financiers et graphes aléatoires

Les réseaux financiers désignent des systèmes financiers interconnectés dans lesquels des échanges d'informations ou des interactions financières ont lieu entre les institutions. Lorsque deux institutions financières sont liées, tout événement financier affectant l'une d'entre elles aura un impact sur son homologue, entraînant des changements dans leurs états financiers respectifs. Ainsi, lorsqu'un réseau financier est confronté à un choc externe, les instabilités peuvent se propager des institutions initialement touchées vers d'autres à travers ces liens interconnectés. Cela peut entraîner des risques importants à l'échelle du système. Le risque systémique fait référence au risque d'une perturbation ou d'un effondrement généralisé et substantiel de l'ensemble d'un système financier ou d'un marché, plutôt que d'une institution ou d'un secteur spécifique. La crise financière de 2007-2009 a illustré l'importance des structures de réseau dans l'amplification des chocs initiaux au sein du système bancaire à un niveau mondial, entraînant une récession économique. Une importante littérature sur le risque systémique et les réseaux financiers a émergé, voir par exemple [91, 146] pour deux études récentes et les références qui y sont mentionnées. En particulier, il a été démontré dans [7, 28, 105, 124, 179] que la topologie du réseau joue un rôle important dans la propagation des défauts dans les systèmes financiers.

Graphes, degrés et structure de connectivité

Nous utiliserons quelques notions de théorie des graphes pour étudier nos réseaux financiers. Commençons par introduire quelques concepts de base. Nous suivons certaines notations et définitions issues de [139], où vous trouverez des informations plus détaillées sur les graphes et les réseaux.

Un graphe $\mathcal{G} = (V, E)$ se compose d'une collection de sommets V, appelée ensemble des sommets, et d'une collection d'arêtes, appelée ensemble des arêtes, E. Les sommets correspondent aux institutions financières que nous modélisons, les arêtes indiquent les interconnexions entre les paires d'institutions. Les graphes peuvent être classés en deux types, non orientés et orientés. Une arête est une paire non ordonnée $u, v \in E$ indiquant que u et v sont directement connectés. Lorsque $\mathcal{G}$ est non orienté, si u est directement connecté à v, alors v est également directement connecté à u. Ainsi, une arête peut être considérée comme une paire de sommets. Dans notre contexte, nous traitons des graphes orientés, où les arêtes sont indiquées par la paire ordonnée (u, v), ce qui signifie une arête allant de u à v. Dans ce cas, lorsque l'arête (u, v) est présente, l'arête inverse (v, u) n'a pas nécessairement besoin d'être présente. Dans un système financier, c'est en réalité le cas, car le créancier et le débiteur jouent des rôles différents dans leur relation mutuelle. Si l'institution u est exposée à l'institution v, alors il y a une arête dirigée de v vers u.

Dans cette thèse, nous considérons de grands réseaux, où l'ensemble des sommets V a une taille importante $n \in \mathbb{N}$. Dans ce cas, nous pouvons numéroter les sommets de 1 à n et supposer que $V = [n] := 1, \ldots, n$, ce que nous ferons à partir de maintenant. Un rôle particulier est joué par le graphe complet, où l'ensemble des arêtes est constitué de toutes les paires possibles de sommets, c'est-à-dire $E = i, j : 1 \leqslant i < j \leqslant n$. Le graphe complet est le graphe le plus fortement connecté sur n sommets, et tout autre graphe peut être considéré comme un sous-graphe de celui-ci, obtenu en conservant certaines arêtes et en supprimant le reste. Les réseaux du monde réel, non limités aux réseaux financiers, présentent une grande diversité dans leur structure de connectivité, tels que les réseaux en anneau, les réseaux en étoile, les réseaux en arbre, etc. Différentes structures de graphes conduisent à des performances différentes dans différents modèles.

Une caractéristique importante des graphes est le degré, qui mesure la connectivité d'un sommet dans le graphe. Dans les graphes non orientés, le degré d_i du sommet $i \in [n]$ est défini comme le nombre d'arêtes contenant i, c'est-à-dire $d_i = \#\{j \in [n] : \{i, j\} \in E\}$. Dans notre contexte, dans les graphes orientés, le degré d se compose de deux parties, le degré d'entrée d^+ et le degré de sortie d^-, qui sont définis, par exemple pour le sommet i,

$$d_i^+ = \#\{j \in [n] : (j, i) \in E\}, \quad \text{et} \quad d_i^- = \#\{j \in [n] : (i, j) \in E\}.$$

Nous appelons toutes les arêtes dirigées vers i $(j, i) \in E : j \in [n]$ les voisins entrants de i, et toutes les arêtes partant de i, c'est-à-dire $(i, j) \in E : j \in [n]$, les voisins sortants de i.

Nous fournirons une analyse quantitative des réseaux financiers selon différents modèles de risque.

Graphe aléatoire avec degrés de sommets donnés

Un point crucial dans la modélisation du risque systémique est la disponibilité de l'information. Si nous disposons de toutes les informations nécessaires sur le réseau financier, y compris sa structure de connectivité, nous pouvons le modéliser et l'analyser efficacement. Cependant, les informations disponibles ne sont pas toujours complètes, en particulier lorsqu'il s'agit de réseaux financiers très vastes. Comme le soulignent [30, 126, 144, 187, 193], seules des informations partielles sont généralement disponibles pour les réseaux financiers, par exemple la taille totale des actifs et des passifs de chaque institution. Pour faire face à une observation incomplète des connexions du système, nous allons utiliser des graphes aléatoires.

Il existe différents types de graphes aléatoires, et parmi eux, nous nous concentrerons sur les graphes avec des degrés fixes. Idéalement, nous aimerions étudier des graphes aléatoires uniformes ayant une séquence de degrés prédéterminée, c'est-à-dire une séquence de degrés qui nous est donnée à l'avance. En général, les informations partielles observées nous permettent de déterminer le nombre de créanciers et de débiteurs pour presque toutes les institutions. Nous analyserons les réseaux financiers en nous basant sur le *Modèle de Configuration*, qui a été initialement développé par Bender et Canfield [57] ainsi que par Bollob'as [65] comme moyen de générer un graphe aléatoire avec une séquence prescrite de degrés de sommets. Ses premières applications étaient dans l'étude des graphes réguliers aléatoires.

Nous étudions le modèle de configuration dirigé. Sans perte de généralité, nous supposons tout au long de cette thèse que $d_i \geqslant 1$ pour tout $i \in [n]$, car lorsque $d_i = 0$, le sommet i est isolé et peut être supprimé du graphe. Étant donné les séquences de degrés $\mathbf{d}^+n = (d_1^+, \ldots, d_n^+)$ et $\mathbf{d}^-n = (d_1^-, \ldots, d_n^-)$ telles que $\sum i \in [n] d_i^+ = \sum i \in [n] d_i^-$, nous associons à chaque institution i deux ensembles : $\mathcal{H}_i^+$ l'ensemble des demi-arêtes entrantes et $\mathcal{H}_i^-$ l'ensemble des demi-arêtes sortantes, avec $|\mathcal{H}_i^+| = d_i^+$ et $|\mathcal{H}_i^-| = d^-i$. Soit $\mathbb{H}^+ = \bigcup i = 1^n \mathcal{H}^+ i$ et $\mathbb{H}^- = \bigcup i = 1^n \mathcal{H}_i^-$. Une *configuration* est un appariement entre $\mathbb{H}^+$ et $\mathbb{H}^-$. Lorsqu'une demi-arête sortante de l'institution i est appariée avec une demi-arête entrante de l'institution j, une arête dirigée de i vers j apparaît dans le graphe. Le modèle de configuration est le multigraphe dirigé aléatoire uniformément distribué sur toutes les configurations. Le graphe aléatoire construit par le modèle de configuration est noté $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$.

Il n'est pas toujours possible de construire un graphe simple avec une séquence de degrés donnée. À la place, nous construisons un multigraphe, qui autorise les boucles et les arêtes multiples entre paires de sommets. Cependant, notre objectif est de modéliser les réseaux financiers comme des graphes aléatoires simples uniformes. Comment abordons-nous cette question ? En réalité, il existe plusieurs approches pour résoudre ce problème. Cependant, nous n'avons pas nécessairement besoin de générer un graphe aléatoire simple uniforme. Grâce à certaines découvertes dans les graphes aléatoires, le modèle de configuration peut offrir des insights et des résultats pour notre objectif. Il est facile de montrer que conditionné à ce que le multigraphe soit un graphe simple, nous obtenons un graphe aléatoire uniformément distribué avec ces séquences de degrés données, noté $\mathcal{G}_*^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$. En particulier, en nous appuyant sur la proposition suivante, toute propriété qui se réalise avec une forte probabilité sur le modèle de configuration se réalise également avec une forte probabilité sur le graphe aléatoire simple uniforme, c'est-à-dire qu'elle se réalise également avec une forte probabilité sur nos réseaux financiers.

Proposition 0.1 ([10])**.** *Toute propriété qui se réalise avec une forte probabilité sur le modèle de configuration se réalise également avec une forte probabilité sur ce graphe aléatoire simple (pour le graphe aléatoire* $\mathcal{G}_*^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$*) à condition que*

$$\liminf_{n\to\infty} \mathbb{P}(\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-) \text{ simple}) > 0.$$

D'autre part, une condition de moment du second ordre sur la séquence de degrés $\sum_{i=1}^n (d_i^+)^2 + (d_i^-)^2 = O(n)$ peut nous garantir que la condition ci-dessus est satisfaite, selon un résultat de Janson [149].

Nous nous référons à [194, Chapitre 7.5] pour plus d'informations sur le modèle de configuration. Comme nous le verrons plus tard, sous certaines restrictions imposées à la séquence de degrés, nos réseaux financiers satisfont toutes les conditions de la proposition mentionnée ci-dessus. Par conséquent, nous pouvons utiliser une approche probabiliste pour étudier diverses propriétés financières dans différents modèles de risque en les associant au modèle de configuration. Dans cette optique, nous pouvons établir nos modèles basés sur le modèle de configuration.

Propagation des défauts et ventes forcées

Pour analyser la contagion des défauts dans les réseaux financiers, il est nécessaire de définir la condition de défaut pour les institutions en fonction des informations disponibles et d'examiner comment la contagion se propage dans les réseaux.

Modèle de seuil et percolation bootstrap

Comme dans [20] et en voir plus de détails dans la partie principale (Partie I), nous utiliserons le nombre de voisins en défaut comme critère. Chaque institution se voit attribuer aléatoirement un seuil de défaut selon une certaine distribution, basée sur ses caractéristiques financières. Lorsque le nombre de voisins en défaut atteint ou dépasse ce seuil de défaut, l'institution elle-même tombe en défaut.

Nous décrirons la dynamique de contagion des réseaux en utilisant la percolation bootstrap. La percolation bootstrap a été introduite par Chalupa, Leath et Reich [86] en 1979 dans le contexte des systèmes magnétiques désordonnés. La percolation bootstrap est un processus de diffusion qui a été étudié sur divers graphes, voir par exemple [9, 10, 148, 150]. Le processus de percolation bootstrap (ainsi que de nombreuses variations de celui-ci) a une longue histoire en physique statistique et est largement utilisé comme modèle pour décrire plusieurs phénomènes complexes dans divers domaines, de la propagation des pandémies [153] à l'activité neuronale [11] et à la propagation des défauts dans les systèmes bancaires [20]. Dans le processus de percolation bootstrap, pour un seuil fixe $\theta \geqslant 2$, il y a initialement un sous-ensemble de nœuds actifs et à chaque tour, chaque nœud inactif ayant au moins θ voisins actifs devient actif et le reste indéfiniment. Récemment, la normalité asymptotique de la percolation bootstrap a également été étudiée dans [13]. La percolation bootstrap est également étroitement liée au problème du k-core dans les graphes aléatoires, car elle se révèle être une méthode puissante pour trouver le k-core, voir par exemple [151].

Dans notre contexte, nous considérons un processus de percolation bootstrap sur le graphe avec une séquence donnée de seuils de défaut. Ce processus est déterministe et évolue par étapes. Chaque institution dans le graphe peut être dans l'un des deux états : solvable ou en défaut (également appelé inactif ou actif dans certaines publications). Initialement, un sous-ensemble de sommets du graphe représente les institutions en défaut, tandis que toutes les autres institutions sont solvables. À chaque étape du processus, si un sommet solvable a un nombre de voisins en défaut supérieur ou égal à son seuil, il devient également en défaut et reste dans cet état de manière permanente. Le processus se poursuit jusqu'à ce qu'aucune autre institution ne devienne infectée, moment où il s'arrête.

Cascade de défauts et résultats asymptotiques

Nous abordons la construction de la contagion des défauts de manière dynamique, où la cascade de défauts évolue étape par étape. La cascade de défauts peut être considérée comme un processus de percolation bootstrap appliqué au modèle de configuration. Si la structure du graphe est connue et que la séquence des seuils est fixée, le processus de percolation devient déterministe. Cependant, en raison de la nature aléatoire du modèle de configuration et de la variabilité de la séquence des seuils, le processus de percolation lui-même devient stochastique, ce qui rend son analyse complexe.

Pour résoudre ce problème, nous proposons de classer toutes les institutions en un ensemble de types dénombrables $\mathcal{X}$ en fonction de leurs caractéristiques financières observées. Les institutions appartenant au même type partagent la même distribution de seuil sur $0, 1, \ldots, d^+$. De plus, nous introduisons un cadre en temps continu pour la contagion de défaut. Une fois qu'une institution fait défaut, elle subit des pertes envers ses voisins sortants après une période de temps stochastique. En conséquence, le processus de contagion de défaut devient un processus stochastique avec des sauts. Notre objectif est d'étudier ses propriétés asymptotiques lorsque la taille du réseau n devient grande. Nous obtenons à la fois des lois des grands nombres (LLN) et des théorèmes centraux limites (TCL). Nous étudions également des processus de cascade de défauts plus généraux dans des réseaux financiers stochastiques et obtenons un résultat de LLN en utilisant une méthode de torsion temporelle pour les processus de Markov, qui apparaît, par exemple, dans l'étude du modèle de pandémie SIR dans [153].

Les résultats présentés dans le chapitre 2 étendent les travaux de [20]. Nous utilisons une approche probabiliste pour établir la loi des grands nombres pour des caractéristiques clés du réseau pendant le processus de contagion de défaut. Ces caractéristiques comprennent (mais ne se limitent pas à) le nombre d'institutions solvables et en défaut tout au long de la dynamique, y compris les quantités finales après l'arrêt de la contagion. Par rapport à [20], nous fournissons des informations plus détaillées sur l'état du réseau pendant la contagion de défaut. Bien que les deux études examinent la cascade de défauts en la construisant comme un processus dynamique, notre preuve diffère de celle de [20], qui repose sur les résultats de fluides limites des équations différentielles. En revanche, nous utilisons une méthode probabiliste. Nous considérons les quantités correspondant aux caractéristiques du réseau comme des processus stochastiques qui évoluent dans le temps. Nous démontrons que ces processus, lorsqu'ils sont normalisés par la taille du réseau, convergent conjointement vers des processus gaussiens. Les fonctions de covariance de ces fluctuations gaussiennes asymptotiques sont données explicitement. En fin de compte, nous prouvons que ces caractéristiques convergent conjointement vers des vecteurs gaussiens après l'arrêt de la contagion. De plus, nous proposons des théorèmes limites pour diverses fonctions d'agrégation de richesse à l'échelle du système et étudions comment le risque systémique peut être lié à l'hétérogénéité des réseaux financiers.

Contributions du Chapitre 2 : Théorèmes de limite pour la contagion de défaut et le risque systémique

Notations. Soit $\{X_n\}_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Si $c \in \mathbb{R}$ est une constante, nous écrivons $X_n \xrightarrow{p} c$ pour indiquer que X_n converge en probabilité vers c, c'est-à-dire que pour tout $\epsilon > 0$, nous avons $\mathbb{P}(|X_n - c| > \epsilon) \to 0$ lorsque $n \to \infty$. Nous écrivons $X_n \xrightarrow{d} X$ pour indiquer que X_n converge en distribution vers X. Soit $\{a_n\}_{n\in\mathbb{N}}$ et $\{b_n\}_{n\in\mathbb{N}}$ deux suites de nombres réels tendant vers l'infini lorsque $n \to \infty$. Nous écrivons $X_n = o_p(a_n)$ si $|X_n|/a_n \xrightarrow{p} 0$, et nous écrivons $X_n = O_p(a_n)$ si $\mathbb{P}(|X_n| \leqslant C|a_n|) \to 1$ lorsque $n \to \infty$ pour une certaine constante C. Nous écrivons $a_n = o(b_n)$ si $a_n/b_n \to 0$, et nous écrivons $a_n = O(b_n)$ si pour une certaine constante C, $|a_n| \leqslant C|b_n|$. Si E_n est un sous-ensemble mesurable de Ω, pour tout $n \in \mathbb{N}$, nous disons que la suite $\{E_n\}_{n\in\mathbb{N}}$ se produit avec une probabilité élevée (w.h.p.) si $\mathbb{P}(E_n) = 1 - o(1)$ lorsque $n \to \infty$. La notation $\mathbb{1}\{E\}$ est utilisée pour l'indicateur d'un événement E ; il vaut 1 si E se produit et 0 sinon.

Nous classifions les institutions financières dans un ensemble de types $\mathcal{X}$, qui est dénombrable. Sans perte de généralité, nous supposons que les institutions appartenant au même type ont le même degré sortant, le même degré entrant et la même distribution des seuils de défaut.

Hypothèses 2.1 et 2.2

- *Il existe une classification des institutions financières dans un ensemble dénombrable de caractéristiques possibles $\mathcal{X}$ tel que, pour chaque $n \in \mathbb{N}$, les institutions de la même classe de caractéristiques ont la même fonction de distribution des seuils (notée $q_x^{(n)}$ pour les institutions de la classe $x \in \mathcal{X}$).*

- *Pour un réseau de taille n, soit $\mu_x^{(n)}$ la distribution des types et $q_x^{(n)}$ la distribution des seuils de défaut du type $x \in \mathcal{X}$. Pour certaines fonctions de distribution de probabilité μ et q sur l'ensemble des caractéristiques $\mathcal{X}$ et indépendantes de n, nous avons $\mu_x^{(n)} \to \mu_x$ et $q_x^{(n)}(\theta) \to q_x(\theta)$ lorsque $n \to \infty$, pour tous les $x \in \mathcal{X}$ et $\theta = 0, 1, \ldots, d_x^+$. De plus, nous supposons que $\sum_{\theta=0}^{d_x^+} q_x(\theta) = 1$ pour tous les $x \in \mathcal{X}$.*

Pour étudier les résultats du TCL, nous devons restreindre notre attention au régime des réseaux clairsemés.

Hypothèses 2.3a-b

- *(a) Nous supposons que, lorsque $n \to \infty$, les degrés moyens convergent et sont finis :*

$$\lambda^{(n)} := \sum_{x \in \mathcal{X}} d_x^+ \mu_x^{(n)} = \sum_{x \in \mathcal{X}} d_x^- \mu_x^{(n)} \longrightarrow \lambda := \sum_{x \in \mathcal{X}} d_x^+ \mu_x \in (0, \infty).$$

- *(b) Nous supposons que pour toute constante $A > 1$, nous avons*

$$\sum_{i=1}^{n} A^{d_i^+} = n \sum_{x \in \mathcal{X}} \mu_x^{(n)} A^{d_x^+} = O(n) \quad \textit{and} \quad \sum_{i=1}^{n} A^{d_i^-} = n \sum_{x \in \mathcal{X}} \mu_x^{(n)} A^{d_x^-} = O(n).$$

Notez que l'hypothèse 2.3b implique 2.3a. Les résultats du TCL ne nécessitent pas l'hypothèse 2.3b. Nous ne mettons en évidence que les résultats du TCL dans cette introduction.

Pour $z \in [0,1]$, nous définissons les fonctions:

$$\begin{aligned} b(d,z,\ell) &:= \mathbb{P}(\mathsf{Bin}(d,z) = \ell) = \binom{d}{\ell} z^{\ell}(1-z)^{d-\ell}, \\ \beta(d,z,\ell) &:= \mathbb{P}(\mathsf{Bin}(d,z) \geqslant \ell) = \sum_{r=\ell}^{d} \binom{d}{r} z^{r}(1-z)^{d-r}, \end{aligned}$$

où $\mathsf{Bin}(d,z)$ désigne la distribution binomiale avec les paramètres d et z. Nous définissons également:

$$\begin{aligned} f_S^{(n)}(z) &:= \sum_{x \in \mathcal{X}} \mu_x^{(n)} \sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta) \beta\big(d_x^+, z, d_x^+ - \theta + 1\big), \quad f_D^{(n)}(z) = 1 - f_S^{(n)}(z), \\ f_W^{(n)}(z) &:= \lambda^{(n)} z - \sum_{x \in \mathcal{X}} \mu_x^{(n)} d_x^- \sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta) \beta\big(d_x^+, z, d_x^+ - \theta + 1\big), \\ s_{x,\theta,\ell}^{(n)}(z) &:= \mu_x^{(n)} q_x^{(n)}(\theta) b\big(d_x^+, 1-z, \ell\big). \end{aligned}$$

Sauf indication contraire, nous définissons toujours les fonctions sans exposant (n) en remplaçant la distribution des seuils $q_x^{(n)}(\theta)$ et la distribution des degrés $\mu_x^{(n)}$ par leur distribution limite $q_x(\theta)$ et μ_x respectivement. Par exemple, nous définissons :

$$s_{x,\theta,\ell}(z) := \mu_x q_x(\theta) b\big(d_x^+, 1-z, \ell\big).$$

Désignons par $D_n(t)$ et $S_n(t)$ respectivement le nombre d'institutions en défaut et le nombre d'institutions solvables à l'instant t pour un réseau de taille n. Pour $x \in \mathcal{X}$, $\theta \in \mathbb{N}$ et $\ell = 0, \ldots, \theta - 1$, nous notons $S_{x,\theta,\ell}^{(n)}(t)$ le nombre d'institutions solvables de type x, avec un seuil de θ et ℓ voisins en défaut à l'instant t. Soit $W_n(t)$ le nombre de demi-arêtes sortantes infectées et $\tau_n^\star$ le temps d'arrêt correspondant à l'arrêt de la contagion de défaut. Remarquez que lorsque le réseau ne contient plus de demi-arêtes sortantes infectées, la contagion s'arrête.

Définissons

$$z^\star := \sup\Big\{ z \in [0,1] : \lambda z - \sum_{x \in \mathcal{X}} \mu_x d_x^- \sum_{\theta=1}^{d_x^+} q_x(\theta) \beta\big(d_x^+, z, d_x^+ - \theta + 1\big) = 0 \Big\},$$

qui peut être considéré comme la limite du temps d'arrêt $\tau_n^\star$.

La normalité asymptotique à la fin du processus de contagion est donnée comme suit.

Théorème 2.7 *Supposons que les hypothèses 3.2 à 3.5b sont satisfaites. Soit* $t^\star = -\ln z^\star$ *et* $\widehat{z}_n$ *le plus grand* $z \in [0,1]$ *tel que* $f_W^{(n)}(z) = 0$. *Si* $z^\star \in (0,1]$ *et* $z^\star$ *est une solution stable, c'est-à-dire* $\alpha := f_W'(z^\star) > 0$, *alors nous avons conjointement*

$$n^{-1/2}(D_n(\tau_n^\star) - nf_D^{(n)}(\widehat{z}_n)) \xrightarrow{d} \mathcal{Z}_D(t^\star) - \alpha^{-1} f_D'(z^\star)\mathcal{Z}_W(t^\star),$$

$$n^{-1/2}(S_n(\tau_n^\star) - nf_S^{(n)}(\widehat{z}_n)) \xrightarrow{d} \mathcal{Z}_S(t^\star) - \alpha^{-1} f_S'(z^\star)\mathcal{Z}_W(t^\star),$$

furthermore, $\widehat{z}_n \xrightarrow{p} z^\star$ *and, for all* $x \in \mathcal{X}$, $0 \leqslant \ell < \theta \leqslant d_x^+$,

$$n^{-1/2}(S_{x,\theta,\ell}^{(n)}(\tau_n^\star) - ns_{x,\theta,\ell}^{(n)}(\widehat{z}_n)) \xrightarrow{d} \mathcal{Z}_{x,\theta,\ell}^*(t^\star) - \alpha^{-1} s_{x,\theta,\ell}'(z^\star)\mathcal{Z}_W(t^\star),$$

où $\mathcal{Z}_D$, $\mathcal{Z}_S$ *et* $\mathcal{Z}_W$ *sont des processus Gaussiens centrés avec des fonctions de covariance caractérisées explicitement.*

Nous étudions également la LCT concernant le risque systémique du réseau financier.

Richesse globale du système : Soit $\bar{\Gamma}_n^\Diamond$ la richesse totale dans le système financier s'il n'y a pas de défaut dans le système. Nous définissons la fonction d'agrégation à l'échelle du système comme suit :

$$\Gamma_n^\Diamond(t) := \bar{\Gamma}_n^\Diamond - \sum_{x\in\mathcal{X}} \bar{L}_x^\odot D_x^{(n)}(t) - \sum_{x\in\mathcal{X}} \bar{L}_x^\Diamond \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell S_{x,\theta,\ell}^{(n)}(t),$$

où $D_x^{(n)}(t)$ est le nombre d'institutions de type x en défaut. Pour chaque type $x \in \mathcal{X}$, nous considérons un coût sociétal fixe borné (dépendant du type) $\bar{L}_x^\odot$ pour les institutions en défaut et un coût fixe borné (dépendant du type d'institutions hôtes) $\bar{L}_x^\Diamond$ pour chaque lien en défaut.

Supposons que $\bar{\Gamma}_n^\Diamond/n \to \bar{\Gamma}^\Diamond$ lorsque la taille du réseau $n \to \infty$. Définissons alors :

$$f_\Diamond^{(n)}(z) := \bar{\Gamma}_n^\Diamond/n - \sum_{x\in\mathcal{X}} \bar{L}_x^\odot f_D^{(n)}(z) - \sum_{x\in\mathcal{X}} \bar{L}_x^\Diamond \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell s_{x,\theta,\ell}^{(n)}(z),$$

$$f_\Diamond(z) := \bar{\Gamma}^\Diamond - \sum_{x\in\mathcal{X}} \bar{L}_x^\odot f_D(z) - \sum_{x\in\mathcal{X}} \bar{L}_x^\Diamond \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell s_{x,\theta,\ell}(z).$$

Le résultat est le suivant.

Théorème 2.8. *Supposons que les hypothèses 3.2 à 3.5a sont satisfaites. Les fonctions d'agrégation finales (à l'échelle du système) satisfont (sous les hypothèses 3.2 à 3.5a) :*

(i) Si $z^\star = 0$, *alors presque toutes les institutions font défaut asymptotiquement pendant la cascade et*

$$\frac{\Gamma_n^\Diamond(\tau_n^\star)}{n} \xrightarrow{p} \bar{\Gamma}^\Diamond - \sum_{x\in\mathcal{X}} \mu_x \bar{L}_x^\odot.$$

(ii) Si $z^\star \in (0,1]$ et $z^\star$ est une solution stable, c'est-à-dire, $f'_W(z^\star) > 0$, alors $\frac{\Gamma_n^\lozenge(\tau_n^\star)}{n} \xrightarrow{p} f_\lozenge(z^\star)$ et, sous l'hypothèse 3.5b,

$$n^{-1/2}\big(\Gamma_n^\lozenge(\tau_n^\star) - n\widehat{f}_\lozenge^{(n)}(\widehat{z}_n)\big) \xrightarrow{d} \mathcal{Z}_\lozenge^\star,$$

où $\mathcal{Z}^\star\lozenge$ est une variable aléatoire gaussienne centrée avec une variance $\sigma^\star\lozenge$ donnée par l'équation (2.62).

Ventes précipitées et prix d'équilibre

Comme mentionné précédemment, lorsqu'une institution en défaut affecte ses voisins sortants, cela peut créer une pénurie de liquidités pour ces voisins. Par conséquent, lorsqu'une institution est exposée à une institution en défaut, elle peut être contrainte de liquider une certaine quantité d'actifs illiquides pour maintenir ses réserves de trésorerie conformément aux contraintes réglementaires. Différents types d'institutions utilisent différentes stratégies de liquidation. Dans le contexte d'une crise financière, des ventes précipitées se produisent lorsqu'une institution tente ou est forcée de vendre une quantité importante d'actifs dans un court laps de temps.

Le chapitre 3 se concentre sur l'étude de l'impact combiné des ventes précipitées et des cascades de défaut sur le risque systémique dans des réseaux financiers complexes lors d'une crise financière. Nous examinons les ventes précipitées instantanées dans les réseaux financiers, en utilisant les résultats du chapitre 2. Le terme "ventes précipitées instantanées" fait référence à un réseau qui réagit rapidement à un choc externe. Dès que le choc se produit, la cascade de défaut et le processus de ventes précipitées associé sont déclenchés simultanément. Contrairement à [106], où les prix des ventes précipitées changent à chaque tour, dans notre étude, les prix des ventes précipitées sont déterminés au début du choc. Les institutions sont contraintes de liquider des quantités aléatoires d'actifs illiquides pour compenser les pertes interbancaires pendant la cascade de défauts. Le processus de contagion dépend désormais des prix de liquidation, car le seuil de défaut est influencé par les caractéristiques financières de l'institution, et la valeur des actifs illiquides en fait partie. Le système financier vise à atteindre un état d'équilibre après la survenue de la contagion de défaut et des ventes précipitées.

Le point clé est de trouver un prix d'équilibre $p_n^\star$ après le choc. Nous adoptons une approche conservatrice et supposons que les actifs illiquides ne peuvent être vendus qu'à un certain prix final. Après toutes les ventes, le marché fixera un prix pour les actifs illiquides en fonction de la fonction de demande inverse g. Soit $\Gamma_n(\tau_n^\star(p); p)$ le montant total vendu à la fin. En utilisant le prix de vente p, le prix donné par la fonction de demande inverse est $g(\Gamma_n(\tau^\star n(p); p)/n)$, où $\tau^\star n(p)$ est le temps d'arrêt final qui dépend également de p. Cela nous conduit à définir le prix d'équilibre de l'actif illiquide comme

$$p_n^\star = \sup\big\{p \in [p_{\min}, p_0] : p \leqslant g(\Gamma_n(\tau_n^\star(p); p)/n)\big\},$$

où p_0 est le prix initial sans ventes forcées et $p$min est le prix minimum lorsque tous les actifs sont vendus. Pour chaque prix de vente fixé $p \in [p\mathrm{min}, p_0]$, nous obtenons les résultats limite, à la fois LLN et TCL, concernant les caractéristiques des réseaux dans le processus combiné de contagion de défaut et de ventes forcées, tels que le processus de vente des parts $\Gamma_n(t; p)$. De plus, sous certaines conditions, nous obtenons également les résultats limite (à la fois LLN et TCL) pour le prix d'équilibre

$p_n^\star$ et pour les caractéristiques liées à la structure du réseau à l'état d'équilibre. Enfin, nous étendons le cadre manipulable à un cadre d'actifs illiquides de types multiples.

Contribution du chapitre 3 : Ventes forcées et cascades de défauts

Nous commençons par formuler une hypothèse sur la fonction de demande inverse g.

Hypothèse 3.1 *Soit $p_{\min} \geqslant 0$. Nous supposons que $g : [0, \gamma_{\max}] \to [p_{\min}, p_0]$ vérifie les conditions suivantes :*

(i) $g(0) = p_0$ *(en l'absence de liquidations, le prix est donné de manière exogène par p_0).*

(ii) $g \in \mathcal{C}^1$ *et est une fonction décroissante de $x \in [0, \gamma_{\max}]$ (le prix décroît avec l'excédent moyen de l'offre x).*

(iii) $g(\gamma_{\max}) = p_{\min} \geqslant 0$.

Hypothèse 3.2 *Il existe une classification des institutions financières en un ensemble dénombrable de classes possibles $\mathcal{X}$ telles que, pour chaque $n \in \mathbb{N}$ et pour tout $p \in [p_{\min}, p_0]$, les institutions appartenant à la même classe ont la même fonction de distribution du seuil (notée $q_x^{(n)}$ pour les institutions de la classe $x \in \mathcal{X}$). Autrement dit, pour tout $i \in [n]$ et tout $\theta \in \mathbb{N}$,*

$$\mathbb{P}(\Theta_i^{(n)}(p) = \theta) = q^{(n)}_{x_i^{(n)}}(\theta; p).$$

Hypothèse 3.3 *Pour certaines fonctions de distribution de probabilité μ et $q(.; p)$ sur l'ensemble des classes $\mathcal{X}$ (indépendamment de n), nous avons $\mu_x^{(n)} \to \mu_x$ et $q_x^{(n)}(\theta; p) \to q_x(\theta; p)$ lorsque $n \to \infty$, pour tous $x \in \mathcal{X}$, $\theta = 0, 1, \ldots, d_x^+$ et $p \in [p_{\min}, p_0]$. Les distributions seuil empiriques satisfont $q_x^{(n)}(\theta; p) \in \mathcal{C}^1$ et $q_x(\theta; p) \in \mathcal{C}^1$ sur $p \in [p_{\min}, p_0]$. De plus, lorsque $n \to \infty$, $\frac{\partial q_x^{(n)}}{\partial p}(\theta; p)$ converge uniformément vers $\frac{\partial q_x}{\partial p}(\theta; p)$ en tant que fonction de p pour tous $x \in \mathcal{X}$ et $\theta = 0, 1, \ldots, d_x^+$.*

Dans le chapitre 3, nous considérons également la possibilité qu'une institution ne fasse jamais défaut, c'est-à-dire qu'elle reste solvable même si toutes ses contreparties font défaut. Nous désignons un tel seuil par ∞. Nous supposons que chaque institution liquide une quantité aléatoire d'actifs illiquides tant qu'elle a un voisin en défaut. Nous supposons que ces liquidations sont i.i.d. et dépendent du type x et du seuil θ. Soit $\bar{\gamma}_x$ la valeur constante de liquidation pour chaque institution initialement en défaut de type x.

Hypothèse 3.4 - Liquidation *La moyenne $\bar{\ell}_{x,\theta}(p)$ et la variance $\varsigma^2_{x,\theta}(p)$ des parts vendues pour chaque liquidation sont toutes deux continues en fonction de p, pour tous les $x \in \mathcal{X}$ et $\theta \in \{0, 1, \ldots, d_x^+\} \cup \{\infty\}$.*

Nous définissons les fonctions suivantes, qui sont les fonctions limites des liquidations,

$$f_{x,\theta}^{(n)}(z;p) := \mu_x^{(n)} q_x^{(n)}(\theta;p)\big(\theta - \sum_{\ell=d_x^+-\theta+1}^{d_x^+} \beta(d_x^+, z, \ell)\big), \quad f_{x,\infty}^{(n)}(z;p) := (1-z)\mu_x^{(n)} q_x^{(n)}(\infty;p) d_x^+,$$

et,

$$f_\Gamma^{(n)}(z;p) := \sum_{x\in\mathcal{X}} \Big(\mu_x^{(n)} \bar{\gamma}_x q_x^{(n)}(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p) f_{x,\theta}^{(n)}(z;p) + \bar{\ell}_{x,\infty}(p) f_{x,\infty}^{(n)}(z;p)\Big).$$

Les versions transformées dans le temps des fonctions ci-dessus sont ensuite définies comme suit,

$$\widehat{f}_{x,\theta}^{(n)}(t;p) := f_{x,\theta}^{(n)}(e^{-t};p), \qquad \widehat{f}_\Gamma^{(n)}(t;p) := f_\Gamma^{(n)}(e^{-t};p),$$

et de même pour les autres fonctions.

Soit $\Gamma_n(t;p)$ le total des parts d'actifs illiquides vendues d'ici le temps t. Nous définissons

$$z_n^\star(p) := \sup\big\{z \in [0,1] : f_W^{(n)}(z;p) = 0\big\},$$

et

$$z^\star(p) := \sup\big\{z \in [0,1] : f_W(z;p) = 0\big\},$$

Nous définissons ensuite $t^\star(p) = -\ln z^\star(p)$ et $t_n^\star(p) = -\ln z_n^\star(p)$.

Soit $f^1(z;p)$ et $f^2(z;p)$ les dérivées partielles par rapport au premier et au deuxième paramètre respectivement. Nous avons le théorème suivant concernant la normalité asymptotique du total final des parts vendues.

Théorème 3.15 *Pour tout p fixé dans l'intervalle $[p_{\min}, p_0]$, lorsque $n \to \infty$, le total final des parts vendues satisfait :*

(i) Sous l'hypothèse 2.3a, si $z^\star(p) = 0$, alors asymptotiquement presque toutes les institutions font défaut après le choc et (lorsque $n \to \infty$)

$$\frac{\Gamma_n(\tau_n^\star;p)}{n} \xrightarrow{p} \sum_{x\in\mathcal{X}} \mu_x \Big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta;p)\Big).$$

(ii) Sous l'hypothèse 2.3b, si $z^\star(p) \in (0,1]$ et que $z^\star(p)$ est une solution stable, c'est-à-dire $\alpha(p) := f_W^1(z^\star(p);p) > 0$, alors

$$n^{-1/2}(\Gamma_n(\tau_n^\star;p) - n f_\Gamma^{(n)}(t_n^\star(p);p)) \xrightarrow{d} \mathcal{Z}_\Gamma(t^\star(p);p) - \alpha(p)^{-1} f_\Gamma^1(z^\star(p);p)\mathcal{Z}_W(t^\star(p);p),$$

où $\mathcal{Z}_\Gamma(t;p)$ et $\mathcal{Z}_W(t;p)$ sont des processus gaussiens dépendant de p.

Nous démontrons également le théorème limite suivant sur le prix donné par la fonction de demande inverse $\kappa_n(p) := g(\Gamma_n(\tau_n^\star(p); p)/n)$.

Théorème 3.17 *Pour tout $p \in [p_{\min}, p_0]$ fixé et lorsque $n \to \infty$, le prix $\kappa_n(p)$ donné par la fonction de demande inverse satisfait :*

(i) Sous l'hypothèse 2.3a, si $z^\star(p) = 0$, alors asymptotiquement presque toutes les institutions font défaut après le choc et (lorsque $n \to \infty$),

$$\kappa_n(p) \xrightarrow{p} g\Big(\sum_{x\in\mathcal{X}} \mu_x\big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta;p)\big)\Big).$$

(ii) Sous l'hypothèse 2.3b, si $z^\star(p) \in (0,1]$ et $z^\star(p)$ est une solution stable, c'est-à-dire $\alpha(p) := f_W^1(z^\star(p);p) > 0$, alors

$$n^{1/2}\Big(\kappa_n(p) - g\big(f_\Gamma^{(n)}(t_n^\star(p);p)\big)\Big) \xrightarrow{d} g'\big(f_\Gamma(z^\star(p);p)\big)\Big[\mathcal{Z}_\Gamma(t^\star(p);p) - \alpha(p)^{-1} f_\Gamma^1(z^\star(p);p)\mathcal{Z}_W(t^\star(p);p)\Big],$$

où g' désigne la première dérivée de g.

Nous obtenons ensuite un théorème limite pour le prix d'équilibre après le choc. Pour le réseau de taille n, nous définissons

$$\bar{p}_n := \sup\big\{p \in [p_{\min}, p_0] : p \leqslant g\big(f_\Gamma^{(n)}(z_n^\star(p);p)\big)\big\}.$$

De manière similaire, définissons son équivalent limite

$$\bar{p} := \sup\big\{p \in [p_{\min}, p_0] : p \leqslant g\big(f_\Gamma(z^\star(p);p)\big)\big\}. \tag{1}$$

Nous disons que $\bar{p}$ est une solution de point fixe *stable* si soit $\bar{p} = p_{\min}$, soit $\bar{p} \in (p_{\min}, p_0]$ et qu'il existe un $\epsilon > 0$ tel que $p < g\big(f_\Gamma(z^\star(p);p)\big)$ pour tout $p \in (\bar{p} - \epsilon, \bar{p})$.

Le résultat concernant le prix d'équilibre est le suivant.

Théorème 3.18 *Lorsque $n \to \infty$, le prix d'équilibre satisfait :*

(i) Sous l'hypothèse 2.3a, si $z^\star(\bar{p}) = 0$ et $\bar{p}$ est une solution stable, alors le prix d'équilibre converge vers $p_n^\star \xrightarrow{p} \bar{p}$, où $\bar{p}$ est la plus grande solution de l'équation à points fixes

$$p = g\Big(\sum_{x\in\mathcal{X}} \mu_x\big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta;p)\big)\Big).$$

(ii) Sous l'hypothèse 2.3b, si $z^\star(\bar{p}) \in (0,1]$ est une solution stable de $f_W(z;\bar{p}) = 0$, c'est-à-dire $\alpha(\bar{p}) := f_W^1(z^\star;\bar{p}) > 0$, et $\bar{p}$ est une solution stable de (1.1), alors

$$n^{1/2}(p_n^\star - \bar{p}_n) \xrightarrow{d} -\rho^{-1}(\bar{p})\mathcal{Z}_V(\bar{p}),$$

où

$$\rho(p) := 1 - g'\big(f_\Gamma(z^\star(p);p)\big)\Big[-f_\Gamma^1(z^\star(p);p)\alpha(p)^{-1}f_W^2(z^\star(p);p) + f_\Gamma^2(z^\star(p);p)\Big],$$

et,

$$\mathcal{Z}_V(p) := -g'\big(f_\Gamma(z^\star;p)\big)\Big[\mathcal{Z}_\Gamma(t^\star(p);p) - \alpha(p)^{-1}f_\Gamma^1(z^\star;p)\mathcal{Z}_W(t^\star(p);p)\Big]$$

est une variable aléatoire gaussienne de moyenne 0.

Processus de risque sur les réseaux

Le chapitre 4 est consacré à l'étude des processus de risque sur les réseaux. Rappelons quelques notions de base sur les processus de risque classiques.

Processus de risque classiques

Le processus de risque classique avec des arrivées de sinistres de type Poisson, également connu sous le nom de modèle de Cramér-Lundberg ([100, 171]), est largement utilisé en gestion quantitative des risques, voir par exemple [103, 173]. Dans ce modèle, le capital agrégé d'un assureur qui démarre avec un capital initial γ, un taux de prime α et des montants de sinistre (L_k) (pertes) est donné par le processus de Poisson composé spectrale négative suivant :

$$C(t) = \gamma + \alpha t - \sum_{k=1}^{\mathcal{N}(t)} L_k,$$

où $L_k, k \in \mathbb{N}$, sont des variables aléatoires non négatives i.i.d. suivant une distribution F avec une moyenne $\bar{L}$, et $\mathcal{N}(t)$ est un processus de Poisson d'intensité $\beta > 0$ indépendant de L_k. Le temps de ruine pour l'assureur avec un capital initial γ est défini par

$$\tau(\gamma) := \inf\{t \mid C(t) \leqslant 0\},$$

(avec la convention que $\inf \varnothing = \infty$) et la question centrale est de trouver la probabilité de ruine

$$\psi(\gamma) := \mathbb{P}(\tau(\gamma) < \infty).$$

Il est connu (voir par exemple [35, 115]) que lorsque $\beta\bar{L} > \alpha$, nous avons $\psi(\gamma) = 1$ pour tout $\gamma \in \mathbb{R}$ et lorsque $\beta\bar{L} < \alpha$, la probabilité de ruine peut être calculée à l'aide de la célèbre formule de Pollaczek-Khinchine donnée par

$$\psi(\gamma) = \left(1 - \frac{\beta\bar{L}}{\alpha}\right)\sum_{k=0}^{\infty}\left(\frac{\beta\bar{L}}{\alpha}\right)^k\left(1 - \widehat{F}^{*k}(\gamma)\right),$$

où

$$\widehat{F}(\gamma) = \frac{1}{\bar{L}}\int_0^\gamma \big(1 - F(u)\big)du,$$

et l'opérateur $(\cdot)^{*k}$ représente la convolution k fois.

Processus de risque en réseau

Des efforts récents ont été consacrés à l'étude des processus de risque sur les réseaux. Dans [56], les auteurs étudient les processus de risque et les probabilités de ruine dans les réseaux bipartites. Cependant, il s'agit plutôt d'une combinaison linéaire de plusieurs processus de risque classiques avec une certaine indépendance. Dans le chapitre 4, nous étudions un modèle de risque plus général sur des réseaux financiers hétérogènes, où les institutions peuvent récupérer du capital au fil du temps, c'est-à-dire qu'il existe une fonction croissante du temps $\alpha_i(t)$ pour chaque agent $i \in [n]$ dans le réseau. Soit $C_i(t)$ le capital total de l'agent i au temps t. Nous considérons le processus de risque stochastique en réseau comme suit, pour chaque agent $i \in [n]$:

$$C_i(t) := \gamma_i(1-\epsilon_i) + \alpha_i(t) - \delta_i - \sum_{j\in[n]:j\to i} L_{ji}\mathbb{1}\{\tau_j + T_{ji} \leqslant t\}, \tag{2}$$

où $\tau_j := \inf\{t : C_j(t) \leqslant 0\}$ désigne le temps de ruine pour l'agent $j \in [n]$ et L_{ji} est la perte interbancaire aléatoire causée par j lorsqu'il fait défaut, γ_i est le montant d'actifs externes exposés au risque, ϵ_i est le choc (fraction perdue des actifs externes), T_{ji} est le délai auquel la perte L_{ji} se produit pour i et δ_i représente la valeur totale des créances détenues par les utilisateurs finaux sur l'agent i (dépôts). Dans le chapitre 2, nous étudions la contagion de défaut sans récupération de capital, où le seuil de défaut dépend uniquement du profil de capital et des pertes interbancaires reçues. Ainsi, dans cette situation, pour chaque institution, le seuil de défaut a une distribution fixe. Mais ici, dans la situation avec récupération de capital, le seuil varie au fil du temps. Nous ne pouvons plus appliquer le modèle de seuil. L'analyse devient plus complexe. C'est également un complément au travail précédent (contagion avec récupération de capital) réalisé dans le chapitre 2. Dans [28], les auteurs étudient la contagion avec récupération dans un cadre similaire, mais avec une récupération sur le seuil et d'une forme spéciale. Ici, nous étudions un cas plus général.

Nous étudions ici la probabilité de ruine pour les processus de risque sur des réseaux à grande échelle. Nous établissons des résultats de LGN pour les structures de réseau en utilisant une approche probabiliste, qui repose sur des connaissances sur la classe de Glivenko-Cantelli et les théorèmes associés. Notre étude englobe différents aspects des processus de risque en réseau et des cas spéciaux déjà abordés dans d'autres travaux. Plus précisément, nous étudions les théorèmes limites liés à la dynamique de contagion et aux probabilités de ruine en réseau pour les processus de risque dans un cadre de réseau stochastique. Nous fournissons également des estimations des probabilités de ruine pour des processus de risque en réseau complexes, qui impliquent à la fois des pertes provenant du réseau et des pertes hétérogènes provenant de sources externes.

Contribution du chapitre 4 : Probabilités de ruine pour les processus de risque dans les réseaux stochastiques

Nous considérons un processus d'intensité de révélation de pertes général, noté $\mathcal{R}_n(t)$, pour décrire l'intensité des révélations de pertes inter-réseaux. Plus précisément, si une perte est révélée au temps $t_1 \in \mathbb{R}_+$, nous attendons un temps exponentiel avec un paramètre $\mathcal{R}_n(t_1)$ jusqu'à la prochaine révélation de perte.

Hypothèse 4.1 *Nous supposons que pour certaines fonctions de distribution de probabilité μ sur $\mathcal{X}$, indépendantes de n, nous avons $\mu_x^{(n)} \to \mu_x$, lorsque $n \to \infty$, pour tout $x \in \mathcal{X}$.*

Hypothèse 4.3 *Nous supposons que la fonction d'intensité de perte $\mathcal{R}_n$ satisfait $\mathcal{R}_n(t) = 0$ pour $t > \tau_n^\star$, et $\mathcal{R}_n(t) = n\mathfrak{R}(t) + o_p(n)$ pour $t \leqslant \tau_n^\star$, où $\mathfrak{R}(t)$ est continue, positive et $\|\mathfrak{R}\|_{L^1} := \int_0^\infty \mathfrak{R}(s)ds < \infty$.*

Nous définissons

$$f_S^{\mathfrak{R}}(t) := \sum_{x\in\mathcal{X}} \mu_x \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t), \quad f_D^{\mathfrak{R}}(t) = 1 - f_S^{\mathfrak{R}}(t),$$

et

$$f_W^{\mathfrak{R}}(t) :=\lambda(1 - \phi^{\mathfrak{R}}(t)) - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t).$$

où $\phi^{\mathfrak{R}}$ et $\mathcal{S}_{x,\theta}^{\mathfrak{R}}$ sont des fonctions caractérisées par les caractéristiques du réseau et $\mathfrak{R}$, et sont définis dans le chapitre 4.

Le résultat principal est le suivant.

Théorème 4.5 *Sous l'hypothèse 2.3a et 3.3, pour toute fonction d'intensité de perte $\mathcal{R}_n$ satisfaisant l'Hypothèse 4.3, nous avons lorsque $n \to \infty$,*

$$\sup_{t\leqslant\tau_n^\star} \Big|\frac{S_{x,\theta}^{(n)}(t)}{n} - \mu_x b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t)\Big| \xrightarrow{p} 0.$$

De plus, lorsque $n \to \infty$,

$$\sup_{t\leqslant\tau_n^\star} \Big|\frac{S^{(n)}(t)}{n} - f_S^{\mathfrak{R}}(t)\Big| \xrightarrow{p} 0, \quad \sup_{t\leqslant\tau_n^\star} \Big|\frac{D^{(n)}(t)}{n} - f_D^{\mathfrak{R}}(t)\Big| \xrightarrow{p} 0,$$

et le processus W_n satisfait

$$\sup_{t\leqslant\tau_n^\star} \Big|\frac{W_n(t)}{n} - f_W^{\mathfrak{R}}(t)\Big| \xrightarrow{p} 0.$$

Définissons

$$t_{\mathfrak{R}}^\star := \inf\{t \in [0,1] : f_W^{\mathfrak{R}}(t) = 0\}.$$

Nous disons que $t_{\mathfrak{R}}^\star < \infty$ est une solution stable de $f_W^{\mathfrak{R}}(t) = 0$ s'il existe un petit $\epsilon > 0$ tel que $f_W^{\mathfrak{R}}(t)$ soit négatif sur l'intervalle $[t_{\mathfrak{R}}^\star, t_{\mathfrak{R}}^\star + \epsilon)$.

Le résultat pour les défauts finaux est le suivant.

Théorème 4.9 *Sous les hypothèses 2.3a et 3.3, et pour toute fonction d'intensité de perte $\mathcal{R}_n$ donnée qui satisfait l'hypothèse 4.3, nous avons lorsque $n \to \infty$:*

(i) Si $\int_0^{t^\star_{\mathfrak{R}}} \mathfrak{R}(s)ds = \lambda$*, alors asymptotiquement tous les agents font faillite à la fin du processus de propagation des pertes, c'est-à-dire :*

$$D^{(n)}(\tau_n^\star) = n - o_p(n).$$

(ii) Si $t^\star_{\mathfrak{R}} < \infty$ *est une solution stable de* $f_W^{\mathfrak{R}}(t) = 0$ *et* $\int_0^{t^\star_{\mathfrak{R}}} \mathfrak{R}(s)ds < \lambda$*, alors la probabilité de ruine d'un agent de type* $x \in \mathcal{X}$ *converge vers :*

$$\frac{D_x^{(n)}(\tau_n^\star)}{n\mu_x^{(n)}} \xrightarrow{p} 1 - \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t^\star_{\mathfrak{R}}),$$

et le nombre total d'agents ruinés satisfait :

$$D^{(n)}(\tau_n^\star) = n \sum_{x\in\mathcal{X}} \mu_x(1 - \sum_{\theta=0}^{d_x^+} \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t^\star_{\mathfrak{R}})) + o_p(n).$$

(iii) Si $t^\star_{\mathfrak{R}} = \infty$ *et* $\|\mathfrak{R}\|_{L^1} < \lambda$*, alors la probabilité de ruine d'un agent de type* $x \in \mathcal{X}$ *converge vers*

$$\frac{D_x^{(n)}(\tau_n^\star)}{n\mu_x^{(n)}} \xrightarrow{p} 1 - \sum_{\theta=0}^{d_x^+} b\big(d_x^+, \|\mathfrak{R}\|_{L^1}/\lambda, \theta\big)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(\infty),$$

et le nombre total d'agents ruinés est donné par :

$$D^{(n)}(\tau_n^\star) = n \sum_{x\in\mathcal{X}} \mu_x(1 - \sum_{\theta=0}^{d_x^+} b\big(d_x^+, \|\mathfrak{R}\|_{L^1}/\lambda, \theta\big)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(\infty)) + o_p(n),$$

où $\mathcal{S}^{\mathfrak{R}}_{x,\theta}(\infty)$ *représente la limite de* $\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)$ *lorsque* t *tend vers l'infini.*

Assumption 4.4 *Nous supposons que, lorsque* $n \to \infty$, $\sum_{i\in[n]}(d_i^+ + d_i^-)^2 = O(n)$.

Les résultats mentionnés ont été obtenus en supposant que la fonction d'intensité de révélation des pertes est connue. Cependant, dans le résultat suivant, nous considérons un cas particulier où l'intensité de révélation des pertes dépend du nombre actuel d'arêtes sortantes infectées non révélées ($W_n(t)$) dans le réseau.

Théorème 4.10 *Soit* $\mathbb{L}_\lambda(\mathbb{R}^+)$ *l'espace de toutes les fonctions continues positives intégrables* f *avec* $\|f\|_1 \leqslant \lambda$*. Supposons que l'intensité de révélation des pertes satisfait* $\mathcal{R}_n(t) = \beta W_n(t)$ *pour une certaine constante* β *et que la séquence de réseaux* $\{\mathcal{G}^{(n)}\}_{n\in\mathbb{N}}$ *satisfait les hypothèses 3.3 et 4.4. Alors nous avons :*

(i) Il existe une solution unique $\mathfrak{R}^\star$ *dans* $\mathbb{L}_\lambda(\mathbb{R}^+)$ *avec une valeur initiale* $\mathfrak{R}^\star(0) = \beta\sum_{x\in\mathcal{X}} \mu_x d_x^-(1 - q_{x,0})$ *à l'équation du point fixe* $\mathfrak{R} = \beta\Psi(\mathfrak{R})$*, où* $\Psi : \mathbb{L}_\lambda(\mathbb{R}^+) \to \mathbb{L}_\lambda(\mathbb{R}^+)$ *est l'application.*

$$\Psi(\mathfrak{R})(t) = \lambda(1 - \phi^{\mathfrak{R}}(t)) - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t).$$

(ii) À mesure que n tend vers l'infini, nous avons

$$\sup_{t\leqslant\tau_n^\star}\Big|\frac{\beta W_n(t)}{n}-\mathfrak{R}^\star(t)\Big|\xrightarrow{p}0,$$

et par conséquent,

$$\sup_{t\leqslant\tau_n^\star}\Big|\frac{S^{(n)}(t)}{n}-f_S^{\mathfrak{R}^\star}(t)\Big|\xrightarrow{p}0\quad et\quad\sup_{t\leqslant\tau_n^\star}\Big|\frac{D^{(n)}(t)}{n}-f_D^{\mathfrak{R}^\star}(t)\Big|\xrightarrow{p}0.$$

Jeux en champ moyen de graphon et systèmes interactifs

Risque systémique et systèmes en champ moyen. Les agents dans les réseaux sont généralement influencés par un groupe d'agents dans ce réseau, qui sont les "voisins" au sens spécifique selon les modèles et les contextes. Parfois, un tel impact peut dépendre de l'ensemble de la population. L'étude du risque systémique et de la contagion de défaut dans les réseaux financiers complexes est de plus en plus liée à la théorie des systèmes en champ moyen et des jeux en champ moyen ces dernières années, voir par exemple [32, 85, 136]. Parmi eux, [85] étudie un modèle d'emprunt et de prêt interbancaire. Dans [32], les auteurs étudient un modèle plus complexe d'emprunt et de prêt intra-et-interbancaire, qui inclut différents groupes de banques, et les impacts financiers proviennent à la fois des banques intergroupe et des banques de différents groupes. Un modèle dynamique en champ moyen pour l'étude du risque systémique et de la cascade de contagion est proposé dans [136]. Les cascades de défaut peuvent être modélisées par un cadre alternatif. Considérez une dynamique de diffusion pour décrire l'évolution du capital de chaque agent. Ensuite, le temps de défaut peut être capturé en utilisant des temps d'atteinte, par exemple le temps d'atteinte à 0 de la diffusion. On peut établir un lien entre la proportion d'agents solvables dans de vastes réseaux financiers et la probabilité de défaut dans l'équation de McKean-Vlasov à mesure que la taille des agents n tend vers l'infini, voir [49, 50, 177]. En fin de compte, les systèmes de particules en champ moyen sont bien adaptés pour modéliser l'évolution des objets d'intérêt dans les réseaux finis, et leurs contreparties limites lorsque $n \to \infty$ peuvent, à leur tour, fournir des informations sur les propriétés ou les comportements des événements financiers qui nous intéressent.

L'étude des systèmes en champ moyen avec des interactions homogènes a une riche histoire, remontant aux travaux de Boltzmann, Vlasov, McKean et d'autres (voir par exemple [33, 154, 172]). On peut les considérer comme des limites des systèmes de particules en interaction, provenant à l'origine de modèles en physique statistique. Des modèles interactifs similaires ont été envisagés pour un large éventail d'applications dans différents domaines, notamment les réseaux bancaires, la biologie, les sciences sociales, etc. (voir par exemple [74, 135, 136]). Les Équations Différentielles Stochastiques Rétrogrades (EDSR) de type champ moyen ont été étudiées précocement dans [72, 73]. De plus, la théorie des jeux en champ moyen, introduite par Lasry et Lions dans [163] et Huang, Caines et Malhamé [141, 142], a suscité une attention considérable au cours des dernières années.

Cependant, une limitation des jeux en champ moyen est l'hypothèse d'homogénéité dans les interactions, qui peut ne pas capturer l'hétérogénéité observée dans les systèmes du monde réel. Afin de prendre en compte l'hétérogénéité des interactions, des systèmes avec des populations multi-types ont été proposés dans de nombreux domaines, voir par exemple [75, 178]. Plus récemment, l'étude des systèmes en champ moyen sur de grands réseaux a attiré une attention croissante, voir [52, 61, 97, 110, 155] et les références qui y sont citées.

Dans de nombreux systèmes du monde réel, y compris les réseaux financiers, l'hétérogénéité est prévalente, car différents participants ont des niveaux variables d'influence les uns sur les autres au sein du système. Cette hétérogénéité découle à la fois de la structure du graphe sous-jacent du système et des caractéristiques diverses des acteurs impliqués. Pour mieux modéliser les interactions hétérogènes dans de tels systèmes, l'étude sur les systèmes interactifs en champ moyen avec des graphes a émergé. Les graphes, introduits par Lovász dans [170], servent de modèles naturels en limite continue pour les graphes grands et denses, offrant un puissant outil pour la modélisation et l'analyse de systèmes complexes avec des interactions hétérogènes. Le concept de systèmes interactifs en champ moyen avec graphes a été proposé et de plus en plus étudié ces dernières années, commençant par la validité et la convergence de grandes populations des systèmes de particules vers les systèmes de graphes limites (voir [47, 60] pour les systèmes directs, et [55] pour les systèmes couplés avant-arrière), jusqu'aux bornes de concentration et la concentration des mesures des systèmes de particules avec graphes (voir [51, 54]).

Outre les diffusions interactives classiques pilotées par des mouvements browniens, les systèmes interactifs pilotés par des mesures aléatoires de Poisson sont également étudiés, par exemple dans [3, 52]. Dans [3], l'auteur étudie les processus Hawkes multivariés sur des graphes hétérogènes et leurs limites de graphes. L'incorporation de la structure de graphe sous-jacente dans la dynamique est étudiée dans [52]. L'utilisation des graphes pour analyser les interactions hétérogènes dans la théorie des jeux en champ moyen est également de plus en plus étudiée, voir [36, 82, 162]. De plus, l'utilisation des graphes pour apprendre les jeux en champ moyen sur des réseaux hétérogènes a récemment émergé, voir par exemple [101, 140]. Dans la deuxième partie de cette thèse, nous nous concentrons d'abord sur l'étude d'un système pur en champ moyen avec un graphe rétrograde et ses mesures de risque associées. Ensuite, nous examinons des problèmes de contrôle stochastique basés sur des systèmes directs en champ moyen avec des sauts.

Graphons. Un graphon est défini comme une fonction mesurable symétrique $G : I \times I \to I$, avec $I = [0,1]$. Les graphon peuvent être considérés comme les limites des matrices d'arêtes de graphes pondérés, lorsque la taille du graphe (nombre de sommets) tend vers l'infini. En effet, en renumérotant les sommets du graphe par i/n, $i \in [n] := \{1,\dots,n\}$, à mesure que n devient grand, les étiquettes $i/n, i \in [n]$ deviennent proches les unes des autres, tendant vers un continuum dans $[0,1]$. Soit $\mathcal{B}(I)$ l'algèbre de Borel sur I. La soi-disant norme de découpe d'un graphon est définie par

$$\|G\|_\square := \sup_{A,B\in\mathcal{B}(I)} \left| \int_{A\times B} G(u,v)dudv \right|.$$

Nous pouvons également considérer un graphe comme un opérateur de $L^\infty(I)$ vers $L^1(I)$, associant à tout $\phi \in L^\infty(I)$:

$$G\phi(u) := \int_I G(u,v)\phi(v)dv.$$

Par Lovász [170, Lemma 8.11], la norme de l'opérateur résultant s'avère équivalente à la norme de découpe

$$\|G\|_{\square} \leqslant \|G\|_{\infty\to 1} \leqslant 4\|G\|_{\square},$$

avec

$$\|G\|_{\infty\to 1} := \sup_{|\phi|\leqslant 1} \|G\phi\|_{L^1}.$$

Ces normes seront utilisées dans l'étude des théorèmes de convergence pour les systèmes de graphon induits par une séquence de graphons. Pour étudier des résultats de convergence plus forts, nous devons considérer une autre norme d'opérateur pour les graphons, en considérant G comme un opérateur de $L^{\infty}(I)$ vers $L^{\infty}(I)$ avec la norme définie par

$$\|G\|_{\infty\to\infty} := \sup_{|\phi|\leqslant 1} \|G\phi\|_{L^\infty}.$$

Avec un espace métrique donné $\mathcal{S}$, nous désignons par $\mathcal{M}_+(\mathcal{S})$ l'ensemble des mesures Borel mesurables non négatives sur $\mathcal{S}$ et par $\mathcal{M}^+_{\mathrm{Unif}}([0,1]\times\mathcal{S})$ l'ensemble des mesures Borel non négatives sur $[0,1]\times\mathcal{S}$ avec une première marge uniforme. Nous définissons la fonction à valeurs de mesure $\Lambda\mu : [0,1] \to \mathcal{M}_+(\mathcal{S})$ pour tout $\mu \in \mathcal{M}^+_{\mathrm{Unif}}([0,1]\times\mathcal{S})$ comme suit:

$$\Lambda\mu(u) := \int_{[0,1]\times\mathcal{S}} G(u,v)\delta_x\mu(dv,dx), \tag{3}$$

où δ_x désigne la mesure de Dirac concentrée en x. Pour toute fonction mesurable bornée $\phi : \mathcal{S} \to \mathbb{R}$, le produit scalaire usuel est défini par

$$\langle \Lambda\mu(u), \phi\rangle = \int_{[0,1]\times\mathcal{S}} G(u,v)\phi(x)\mu(dv,dx).$$

Risque systémique sur de grands réseaux hétérogènes. Dans la Partie I, bien que nous classifiions les institutions financières par le biais d'un ensemble de caractéristiques $\mathcal{X}$, le réseau présente une probabilité de connexion égale entre les institutions de différents types. Une extension significative consisterait à introduire une probabilité de connexion hétérogène entre les institutions, qui peut être modélisée par une distribution de choix, disons $Q_x(\cdot)$ sur l'ensemble $\mathcal{X}$ pour chaque type $x \in \mathcal{X}$, c'est-à-dire que pour chaque opportunité de connexion, une institution de type x tend à choisir une institution de type y avec une probabilité $Q_x(y)$ de manière indépendante. Ensuite, la probabilité de connexion entre une paire de types (x,y) est $Q_x(y)Q_y(x)$, ce qui peut être réécrit comme une fonction symétrique $\bar{Q}(x,y)$ sur $\mathcal{X}^2$. Ici, la fonction $\bar{Q}(x,y)$ joue un rôle similaire au graphon G, montrant la pertinence d'introduire des graphons dans l'étude du risque systémique dans de grands réseaux hétérogènes. Dans un tel cadre, l'étude de la percolation dans un graphe hétérogène est intéressante. Dans [48], les auteurs étudient le problème du k-core dans des séquences de graphes hétérogènes denses percolés convergents au sens de la norme de découpe. Cela pourrait être lié à l'étude des modèles de risque dans de grands réseaux hétérogènes. Les travaux futurs pourraient inclure des extensions des modèles de la Partie I impliquant des graphons. L'étude des jeux entre les institutions (par exemple, la connectivité optimale, la probabilité de connexion optimale) ou des problèmes d'optimisation impliquant un régulateur extérieur (interventions ciblées) sont également des sujets intéressants à étudier. Dans la Partie II, nous nous concentrons sur les systèmes et les jeux de champ moyen de graphon. L'étude du risque systémique dans des modèles de champ moyen de graphon complexes est réservée aux travaux futurs.

Équations différentielles stochastiques rétrogrades à sauts en champ moyen de graphon

Dans le Chapitre 5, nous étudions les équations différentielles stochastiques rétrogrades à sauts en champ moyen de graphon et les mesures de risque dynamiques associées. Nous considérons un système rétrograde. Les systèmes en champ moyen de graphon avant et avant-arrière sans sauts ont été étudiés respectivement dans [47] et [55]. Nous considérons l'équation BSDE en champ moyen de graphon suivante avec sauts :

$$\begin{aligned} X_u(t) = \xi_u &+ \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx)dyds - \int_t^T Z_u(s)dW_u(s) \\ &- \int_t^T \int_E \ell_{u,s}(e)\widetilde{N}_u(ds,de), \qquad\qquad u \in I, \quad \text{pour} \quad t \in [0,T], \end{aligned} \tag{4}$$

où $\mu_y := \mathcal{L}(X_y) \in \mathcal{P}(\mathcal{D})$ et $\mu_{y,s} := \mathcal{L}(X_y(s)) \in \mathcal{P}(\mathbb{R})$. Nous supposons que pour chaque $u \in I$, $\xi_u \in L^2(\mathcal{F}_T)$ et que l'application $u \mapsto \xi_u$ est mesurable.

L'interaction hétérogène est gouvernée par le terme du graphon G. Notez que si l'interaction est homogène, alors $G(u,v) \equiv 1$ pour tous $(u,v) \in [0,1] \times [0,1]$. Dans ce cas, l'équation BSDE avec sauts ci-dessus se réduit au cas standard en champ moyen,

$$\begin{aligned} X_u(t) = \xi_u &+ \int_t^T \int_I \int_{\mathbb{R}} f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx)dyds - \int_t^T Z_u(s)dW_u(s) \\ &- \int_t^T \int_E \ell_{u,s}(e)\widetilde{N}_u(ds,de), \qquad\qquad u \in I, \quad \text{pour} \quad t \in [0,T]. \end{aligned}$$

Ce cas particulier a été étudié dans [89]. Dans notre recherche, nous menons une analyse approfondie des BSDE en champ moyen de graphon avec sauts. Nous établissons certains résultats fondamentaux, notamment l'existence et l'unicité des solutions, des estimations pour les solutions et des théorèmes de comparaison. De plus, nous explorons également la propagation du chaos de ses systèmes de particules N associés. Plus précisément, nous considérons un système de BSDE couplé à N, où chaque équation est indexée par $i = 1, \ldots, N$, et a la forme suivante :

$$
\begin{aligned}
X_i^N(t) &= \xi_i^N + \int_t^T \frac{1}{N}\sum_{j=1}^N \zeta_{ij}^N f(s, X_j^N(s), X_i^N(s), Z_i^N(s), \ell_s^{N,i}(\cdot))ds - \int_t^T Z_i^N(s)d\widehat{W}_i(s) \\
&\quad - \int_t^T \int_E \ell_s^{N,i}(e)\widetilde{\widehat{N}}_i(ds, de), \quad t \in [0,T] \\
X_i^N(T) &= \xi_i^N,
\end{aligned}
\tag{5}
$$

où $\widehat{W}_i := W_{\frac{i}{N}}$ sont des mouvements browniens i.i.d., et $\widehat{N}_i(dt, de) = N_{\frac{i}{N}}(dt, de)$ sont des mesures aléatoires de Poisson indépendantes. Nous supposons que $\xi_i^N \in L^2(\mathcal{F}_T)$ pour tous les $i = 1, \ldots, N$. Ici, $\zeta_{ij}^N : \mathbb{N} \times \mathbb{N} \mapsto \mathbb{R}_0^+$ est symétrique décrivant la force d'interaction entre la particule i et j.

Le graphon G peut être considéré comme la limite de ζ_{ij}^N lorsque $N \to \infty$. Nous étudions les deux types de convergence suivants pour la solution :

- Le type moyenne :

$$
\frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right].
$$

- Le type maximum :

$$
\max_{i\in[N]} \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right].
$$

Nous introduisons les mesures de risque dynamiques en champ moyen de graphon induites par les solutions des BSDE en champ moyen de graphon avec sauts. Nous étendons plusieurs propriétés des mesures de risque dynamiques au cas en champ moyen de graphon, notamment la cohérence, la continuité, l'homogénéité, l'invariance par translation, la monotonie, la convexité et l'absence d'arbitrage. Ces propriétés ont été précédemment étudiées dans le contexte des mesures de risque dynamiques sans terme en champ moyen dans [183] et avec terme en champ moyen dans [89]. De plus, nous fournissons une formule de représentation duale, un résultat fondamental dans la théorie des mesures de risque convexes. Cette formule fournit une méthodologie pour le calcul des mesures de risque dynamiques en prenant le supremum sur un ensemble d'attentes sous une famille de mesures de probabilité. À travers une démonstration complexe, nous établissons la formule de représentation duale pour les mesures de risque dynamiques en champ moyen de graphon.

Jeux stochastiques avec interactions de champ moyen de graphon

Avec un intérêt croissant pour les systèmes interagissant avec des graphes, il y a eu une activité de recherche croissante sur les jeux de champ moyen de graphon. Étudier les jeux stochastiques avec des interactions hétérogènes ou les jeux sur des réseaux pose des défis substantiels, en particulier lorsqu'il s'agit de jeux sur des réseaux impliquant un grand nombre de joueurs, car ces jeux peuvent présenter une asymétrie significative. Cette distinction est particulièrement notable dans le contexte des jeux sur des réseaux clairsemés (par exemple, [120, 161]). L'analyse des jeux sur de grands réseaux, en particulier ceux avec des interactions hétérogènes, repose souvent sur des modèles limites (continus) traitables. Ces modèles fournissent une approximation pratique pour comprendre la dynamique des jeux finis de grande envergure et offrent des informations précieuses sur les complexités des jeux stochastiques dans les systèmes interagissant de manière hétérogène.

Le Chapitre 6 vise à développer un modèle d'interaction de graphon pour résoudre les jeux en graphes avec des interactions hétérogènes et des sauts, tout en maintenant une traitabilité comparable aux jeux de champ moyen traditionnels (MFG). Le cadre MFG traditionnel repose sur un problème de point fixe décrivant la loi du processus d'état $(X(t))_{t\in[0,T]}$ d'un joueur typique. Dans le modèle de jeu de graphon, nous considérons un problème de point fixe pour une famille de lois $(X_u(.))_{u\in I}$, qui peut être vue comme une loi conjointe de (U, X), où X est le processus d'état aléatoire et la variable aléatoire uniforme U dans $I := [0,1]$ est interprétée comme la variable "étiquette" (ordre du sommet sur le réseau dans un sens limite) du joueur dans le graphique. Malgré les interactions hétérogènes, nous incluons également des sauts dans la dynamique pour modéliser les impacts instantanés. Les sauts sont induits par des mesures aléatoires de Poisson avec des mesures d'intensité différentes pour différentes étiquettes, ce qui est une source d'hétérogénéité individuelle.

Nous portons notre attention sur les contrôles de rétroaction markoviens. Le contrôle dépend de l'état actuel et de son étiquette. Soit $\mathcal{A}_I$ l'ensemble des contrôles de graphon α définis comme une fonction mesurable $\alpha : [0,T] \times I \times \mathbb{R} \to A; (t,u,x) \mapsto \alpha(t,u,x)$, où A est l'ensemble des actions. La dynamique du système de graphon contrôlé est la suivante :

$$
\begin{aligned}
dX_u^\alpha(s) = &\int_I \int_{\mathbb{R}} G(u,v) b(s, X_u^\alpha(s), x, \alpha(s,u,X_u^\alpha(s)))\mu_{v,t}^\alpha(dx)dvds \\
&+ \int_I \int_{\mathbb{R}} G(u,v)\sigma(s, X_u^\alpha(s), x, \alpha(s,u,X_u^\alpha(s)))\mu_{v,t}^\alpha(dx)dvdW_u(s) \\
&+ \int_E \ell(s, X_u^\alpha(s), e, \alpha(s,u,X_u^\alpha(s)))\widetilde{N}_u(ds,de), \qquad X_u(0) = \xi_u, \qquad u \in I,
\end{aligned} \tag{6}
$$

où $\mu_v^\alpha := \mathcal{L}(X_v^\alpha) \in \mathcal{P}(\mathcal{D})$ et $\mu_{v,s}^\alpha := \mathcal{L}(X_v^\alpha(s)) \in \mathcal{P}(\mathbb{R})$. Nous supposons que $\boldsymbol{\xi} := \{\xi_u\}_{u\in I} \in \mathcal{M}L^2(\mathcal{F}_0)$, c'est-à-dire que pour chaque $u \in I$, $\xi_u \in L^2(\mathcal{F}_T)$ et que l'application $u \mapsto \xi_u$ est mesurable. Les coefficients $b : [0,T] \times \mathbb{R} \times \mathbb{R} \times A \to \mathbb{R}$, $\sigma : [0,T] \times \mathbb{R} \times \mathbb{R} \times A \to \mathbb{R}$ et $\ell : [0,T] \times \mathbb{R} \times E \times A \to \mathbb{R}$ sont Lipschitz continus par rapport à tous les paramètres sauf t. Nous supposons également que σ^2 est borné de 0. Noticez que dans notre modèle, le terme de contrôle est présent non seulement dans la dérive, comme dans [82, 162], mais est également présent dans les termes de diffusion et de sauts. De plus, nous avons également l'interaction de graphon dans le terme de diffusion, qui n'est pas présent dans le modèle de [82, 162]. Combinés avec les sauts et les contrôles, plus d'hétérogénéité est introduite dans notre configuration, et le système dynamique interactif devient plus complexe par rapport à [47, 55]. Chaque joueur avec l'étiquette $u \in I$ cherche à maximiser la fonction objective suivante :

$$\mathbb{E}\Big[\int_0^T f(t, X_u^\alpha(t), \Lambda\mu_t^\alpha(u), \alpha(t, u, X_u^\alpha(t)))dt + g(X_u^\alpha(T), \Lambda\mu_T^\alpha(u))\Big],$$

où f est une fonction représentant le coût en cours d'exécution et g est la fonction de coût à l'instant final.

Équilibres de graphon et équilibres de Nash approximatifs

L'étude des jeux de champ moyen de graphon peut aider à étudier les jeux finis sur de grands réseaux. Il est difficile d'étudier directement les équilibres de Nash des jeux finis avec une interaction hétérogène. Au lieu de cela, nous l'étudions à travers les jeux de graphon. Soit $\mathcal{A}_n$ l'ensemble des fonctions mesurables $\alpha : [0, T] \times \mathbb{R}^n \to A$. Le système de particules interagissant de manière hétérogène que nous considérons a la dynamique contrôlée suivante sous le contrôle $\{\alpha_i\}_{i\in[n]} \in \mathcal{A}_n^n$,

$$\begin{aligned} dX_i^{(n)}(s) =& \frac{1}{n}\sum_{j=1}^n \zeta_{ij}^{(n)} b(s, X_i^{(n)}(s), X_j^{(n)}(s), \alpha_i(s, X^{(n)}(s)))ds \\ &+ \frac{1}{n}\sum_{j=1}^n \zeta_{ij}^{(n)} \sigma(s, X_i^{(n)}(s), X_j^{(n)}(s), \alpha_i(s, X^{(n)}(s)))dW_i(s) \\ &+ \int_E \ell(s, X_i^{(n)}(s), e, \alpha_i(s, X^{(n)}(s)))\tilde{N}_i(ds, de), \qquad X_i^{(n)}(0) = \xi_i^{(n)}, \end{aligned} \tag{7}$$

où $\{W_i, i \in [n]\}$ sont des mouvements browniens i.i.d., $\{N_i(dt, de), i \in [n]\}$ sont des mesures aléatoires de Poisson indépendantes, et $\{\xi_i^{(n)}, i \in [n]\}$ sont les conditions initiales. L'idée est que lorsque la taille de la population n est suffisamment grande, le système de champ moyen de graphon contrôlé (1.7) peut être considéré comme le système limite du système couplé contrôlé n (1.8) d'une certaine manière. Ainsi, l'équilibre de Nash du jeu fini devrait être proche de l'équilibre du jeu de graphon, et par conséquent pourrait être approché par celui du jeu de graphon. Nous appelons un tel équilibre du jeu de graphon l'équilibre de graphon et en donnerons la définition détaillée. Des travaux récents [36, 82, 162] ont étudié de tels équilibres de Nash approximatifs dans certains cas spéciaux. Dans le Chapitre 6, nous étendons l'étude à un cadre plus général impliquant des sauts.

Nous définissons le concept d'équilibre de graphon dans notre cadre et en établissons l'existence en utilisant la méthode de la compacification, une technique puissante couramment utilisée pour étudier les équilibres dans divers types de jeux de champ moyen. Nous enquêtons également sur l'unicité de l'équilibre de graphon sous certaines conditions de monotonie. En considérant les équilibres de graphon comme des points de référence, nous pouvons approximer les équilibres de Nash dans les jeux finis. Nous utilisons le contrôle d'équilibre pour les jeux de graphon comme point de référence pour déduire le contrôle correspondant pour les jeux finis. En utilisant les résultats de propagation du chaos (similaires à ceux du Chapitre 5), à mesure que la taille de la population augmente, les distributions des processus d'état dans les jeux finis convergent vers celles des jeux de graphon selon un schéma de correspondance spécifique entre l'ordre du joueur $i \in [n]$ et l'étiquette du graphon $u \in I$. De manière intuitive, le contrôle d'équilibre pour chaque joueur dans le jeu fini devrait ressembler étroitement à celui pris pour l'étiquette correspondante dans le système de graphon limite. Par conséquent, il est naturel de sélectionner le contrôle associé à l'étiquette $\frac{i}{n}$ pour le joueur i-ème dans un jeu à n joueurs. Lorsque le contrôle d'équilibre pour le jeu de graphon présente une continuité par rapport à u, nous pouvons nous détendre pour considérer des contrôles associés à des étiquettes proches de $\frac{i}{n}$. Cette approximation suit les principes de la théorie classique des jeux de champ moyen. Cependant, en raison de l'hétérogénéité des interactions dans notre modèle, l'analyse devient plus complexe. Cette méthode d'approximation pour l'équilibre de Nash s'applique non seulement au cas dépendant du modèle, mais aussi au cas indépendant du modèle. Elle a également trouvé des applications dans l'apprentissage par renforcement, comme on peut le voir dans des travaux tels que [101, 134].

Chapter 1

Introduction

This thesis is divided in two parts. The first part considers the issues of stability and systemic risk in large complex financial networks, including the study of default contagion, fire sales and risk processes on networks. We first prove limit theorems (law of large numbers and central limit theorem types) for the contagion dynamics. We show how to quantify the systemic risk for a financial network under partial information facing an outside shock. Then we present a general tractable framework for understanding the joint impact of fire sales and default cascades on systemic risk in complex financial networks. We finally study risk processes on large financial systems, when agents, located on a large network, receive losses from their neighbors.

The second part of the thesis focuses on graphon mean field interacting systems with jumps and graphon mean field games. Here, the financial network is seen as a large interacting system, with a graphon mean field structure depending on the underlying graph structure of the network. We first conduct a comprehensive study of graphon mean field backward stochastic differential equations (BSDEs) with jumps and associated global dynamic risk measures. We then study continuous stochastic games with heterogeneous mean field interactions on large networks and investigate their graphon limits. We provide approximate Nash equilibria for finite games with heterogeneous interactions, using their graphon equilibria as benchmarks.

1.1 Financial networks and random graphs

Financial networks refer to interconnected financial systems when there are information exchanges or financial interactions between institutions. When two financial institutions are linked, any financial event impacting one of them will affect its counterpart, resulting in changes to their financial states. Therefore, when a financial network faces an external shock, instabilities can propagate from the initially affected institutions to others through these interconnected links. This can give rise to significant risks at the system level. Systemic risk refers to the risk of a widespread and substantial disruption or collapse of an entire financial system or market, rather than just a specific institution or sector. The financial crisis of 2007-2009 illustrated the importance of network structures in amplifying initial shocks within the banking system to a global level, leading to an economic recession. An important literature on systemic risk and financial networks has emerged, see e.g. [91, 146] for two recent surveys and references therein. In particular, it is shown in [7, 28, 105, 124, 179] that network topology plays an important role for default propagation in financial systems.

As the world's connection becomes more and more compact, network structures are becoming increasingly complex. It is thus significant to develop mathematical models to study large networks with complex (random) structures. In particular, networks do not in general appear with observable and fixed size. In this context, limit theorems can be useful to get insights in modeling and monitoring various contagion effects in large financial networks. The first part of this thesis focuses on studying limit theorems in different risk models within large financial networks, encompassing results like the law of large numbers and central limit theorem. These results can provide information regarding systemic risk in networks, considering basic parameters and observable data. Feasibility and stability studies are also included. Various analyses could then be conducted, such as quantifying systemic risk, targeting interventions or optimizing investments.

1.1.1 Graphs, degrees and connectivity structure

We shall use some techniques of graph theory to study our financial networks. Let us first introduce some basic concepts. We follow certain notations and definitions from [139], where more in-depth information about graphs and networks can be found.

A graph $\mathcal{G} = (V, E)$ consists of a collection of vertices V, called the vertex set, and a collection of edges, called the edge set, E. The vertices correspond to the financial institutions that we model, the edges indicate the interlinkages between pairs of institutions. Graphs can be classified into two types, undirected and directed. An edge is an unordered pair $\{u, v\} \in E$ indicating that u and v are directly connected. When $\mathcal{G}$ is undirected, if u is directly connected to v, then v is also directly connected to u. Thus, an edge can be seen as a pair of vertices. In our setting, we deal with directed graphs, where edges are indicated by the ordered pair (u, v), which means an edge with direction from u to v. In this case, when the edge (u, v) is present, the reverse edge (v, u) need not be present necessarily. In a financial system, this is actually the case, since the creditor and debtor take different roles in their counterpart relation. If institution u is exposed to institution v, then there is a directed edge from v to u.

In this thesis, we consider large networks, where the vertex set V has a large size $n \in \mathbb{N}$. In this case, we can number the vertices as $1, 2, ..., n$ and assume that $V = [n] := \{1, \ldots, n\}$, which we will do from now on. A special role is played by the complete graph, where the edge set consists of all possible pairs of vertices, i.e., $E = \{\{i, j\} : 1 \leqslant i < j \leqslant n\}$. The complete graph is the most highly connected graph on n vertices, and every other graph can be considered as a subgraph of it obtained by keeping some edges and removing the rest. The networks in real world, not restricted to financial networks, exhibit a huge diversity in the connectivity structure, such as the ring networks, star networks, tree networks, etc. Different graph structures lead to different performances in various models. An important characteristic in graphs is the degree, which measures the connectivity of a vertex in the graph. In undirected graphs, the degree d_i of vertex $i \in [n]$ is defined as the number of edges containing i, i.e., $d_i = \#\{j \in [n] : \{i, j\} \in E\}$. For our purpose, in directed graphs, the degree d contains two parts, the in degree d^+ and out degree d^-, which are defined as, say for vertex i,

$$d_i^+ = \#\{j \in [n] : (j, i) \in E\}, \quad \text{and} \quad d_i^- = \#\{j \in [n] : (i, j) \in E\}.$$

We call all vertices with edges to i, i.e. the set $\{(j, i) \in E : j \in [n]\}$ incoming neighbors of i, and all vertices with edges departing from i, i.e. the set $\{(i, j) \in E : j \in [n]\}$, outgoing neighbors of i.

We shall provide quantitative analysis of financial networks under different risk models.

1.1.2 Random graph with given vertex degrees

A crucial point in systemic risk modeling is the availability of information. If all the necessary information for the financial network, including the connectivity structure is known, we can effectively model and analyze it. However, the available information is not always complete, particularly when dealing with very large financial networks. As pointed out in [30, 126, 144, 187, 193], only partial information is, in general, available for the financial networks, e.g. the total size of the assets and liabilities for each institution. To deal with an incomplete observation of the system connections, we will take advantage of random graphs.

There are various types of random graphs, and among them, we will focus on graphs with fixed degrees. We investigate uniform random graphs that have a predetermined degree sequence, meaning a degree sequence given to us in advance. Typically, the observed partial information allows us to determine the number of creditors and debtors for almost all institutions. We will analyze the financial networks based on the *Configuration Model*, which was originally developed by Bender and Canfield [57] and Bollobás [65] as a mean for generating a random graph with a prescribed sequence of vertex degrees. Its earliest applications were in the study of random regular graphs.

We consider the directed configuration model. Without loss of generality, we assume throughout this thesis that $d_i \geqslant 1$ for all $i \in [n]$, since when $d_i = 0$, vertex i is isolated and can be removed from the graph. Given the degree sequences $\mathbf{d}_n^+ = (d_1^+, \ldots, d_n^+)$ and $\mathbf{d}_n^- = (d_1^-, \ldots, d_n^-)$ such that $\sum_{i \in [n]} d_i^+ = \sum_{i \in [n]} d_i^-$, we associate to each institution i two sets: $\mathcal{H}_i^+$ the set of in half-edges and $\mathcal{H}_i^-$ the set of out half-edges, with $|\mathcal{H}_i^+| = d_i^+$ and $|\mathcal{H}_i^-| = d_i^-$. Let $\mathbb{H}^+ = \bigcup_{i=1}^n \mathcal{H}_i^+$ and $\mathbb{H}^- = \bigcup_{i=1}^n \mathcal{H}_i^-$. A *configuration* is a matching of $\mathbb{H}^+$ with $\mathbb{H}^-$. When an out half-edge of institution i is matched with an in half-edge of institution j, a directed edge from i to j appears in the graph. The configuration model is the random directed multigraph that is uniformly distributed across all configurations. The random graph constructed by the configuration model is denoted by $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$.

Note that it is not always feasible to construct a simple graph with a given degree sequence. Instead, we construct a multigraph, which allows for self-loops and multiple edges between vertex pairs. Our aim is to model financial networks as uniform simple random graphs. There are several approaches to tackle this problem. We do not necessarily need to generate a uniform simple random graph. The configuration model can offer insights and results for our purpose. It is easy to show that conditioned on the multigraph being a simple graph, we obtain a uniformly distributed random graph with these given degree sequences denoted by $\mathcal{G}_*^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$. In particular, relying on the following proposition, any property which holds with high probability on the configuration model also holds with high probability on the uniform simple random graph, i.e., it also holds with high probability on our financial networks.

Proposition 1.1 ([10])**.** *Any event which holds with high probability on the configuration model also holds with high probability on this random graph being simple (for the random graph $\mathcal{G}_*^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$) provided that*

$$\liminf_{n\to\infty} \mathbb{P}(\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-) \textit{ simple}) > 0.$$

On the other hand, a second moment condition on the degree sequence $\sum_{i=1}^n (d_i^+)^2 + (d_i^-)^2 = O(n)$ guarantees that the above condition holds according to a result by Janson [149].

We refer to [194, Chapter 7.5] for more information on the configuration model. As we will see later, under certain restrictions imposed on the degree sequence, our financial networks satisfy all the conditions in the aforementioned proposition. Consequently, we can use a probabilistic approach to study various financial properties in different risk models by mapping them to the configuration model. With this in mind, we can establish our models based on the configuration model.

1.2 Default cascade and fire sales

To analyze the default contagion in financial networks, it is necessary to define the default condition for institutions based on the available information and examine how the contagion spreads throughout the networks.

1.2.1 Threshold model and bootstrap percolation

As in [20], we will use the number of defaulted neighbors as a criterion. Each institution is randomly assigned a default threshold according to certain distribution, which is based on its financial characteristics. When the number of defaulted neighbors reaches or exceeds this default threshold, the institution itself defaults.

We describe the contagion dynamics of the networks by using bootstrap percolation. Bootstrap percolation was introduced by Chalupa, Leath and Reich [86] in 1979 in the context of magnetic disordered systems. Bootstrap percolation is a diffusion process that has been studied on a variety of graphs, see e.g., [9, 10, 148, 150]. This process (as well as numerous variations of it) has a rich history in statistical physics and in modelling complex phenomena in a diversity of areas, from pandemic spread [153] to neuronal activity [11] and spread of defaults in banking systems [20]. In a bootstrap percolation process, for a fixed threshold $\theta \geqslant 2$, there is an initially subset of active nodes and in each round, each inactive node that has at least θ active neighbors becomes active and remains so forever. Recently, the asymptotic normality of bootstrap percolations has also been studied in [13]. Bootstrap percolation is also closely related to the k-core problem in random graphs, as it shows to be a powerful method to find the k-core, see e.g. [151].

In our context, we consider a bootstrap percolation process on the graph with given defaulted threshold sequence. This process is deterministic and evolves in rounds. Each institution in the graph can be in one of two states: solvent or defaulted (also referred to as active or inactive in some literature). Initially, a subset of vertices in the graph represents defaulted institutions, while all other institutions are solvent. During each round of the process, if a solvent vertex has a number of defaulted neighbors that is greater than or equal to its threshold, it also defaults and remains in this state permanently. The process continues until no more institutions become infected, at which point it stops.

1.2.2 Default cascade and asymptotic results

We construct the default contagion in a dynamic manner, when the default cascade evolves round by round. The default cascade can be seen as a bootstrap percolation process applied to the configuration model. If the graph structure is known and the threshold sequence is fixed, the percolation process becomes deterministic. However, due to the random nature of the configuration model and the variability of the threshold sequence, the percolation process itself becomes stochastic, making it challenging to analyze.

To address this, we propose classifying all institutions into a countable type set $\mathcal{X}$ based on their observed financial characteristics. Institutions belonging to the same type share the same threshold distribution over $\{0, 1, \ldots, d^+\}$. Additionally, we introduce a continuous-time framework for the default contagion. Once an institution defaults, it incurs losses to its outgoing neighbors after a stochastic period of time. As a result, the default contagion process becomes a stochastic process with jumps. Our focus lies in studying its asymptotic properties as the network size n becomes large. We obtain both law of large numbers (LLN) and central limit theorems (CLT). We also study more general default cascade processes in stochastic financial networks and obtain LLN result by employing the time twist method for Markov processes, which appears e.g. in the study of pandemic models in [153].

The results presented in Chapter 2 extend the work of [20]. We use a probabilistic approach to establish the law of large numbers for key network features during the default contagion process. These features contain (but are not limited to) the number of solvent institutions and defaulted institutions throughout the entire dynamics, including the final quantities after the contagion stops. Compared to [20], we provide more detailed information about the state of the network during the default contagion. While both studies investigate the default cascade by constructing it as a dynamical process, our proof differs from that of [20], which relies on limit fluid results of differential equations. In contrast, we utilise a probabilistic method. We consider the quantities that correspond to network features as stochastic processes that evolve over time. We demonstrate that these processes, when normalised by the network size, converge jointly to some Gaussian processes. The covariance functions of these asymptotic Gaussian fluctuations are given explicitly. Ultimately, we prove that these features converge jointly to Gaussian vectors after the contagion stops. Additionally, we provide limit theorems for various system-wide wealth aggregation functions and study how systemic risk can be linked to the heterogeneity of financial networks.

Notation. We first introduce some notation. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If $c \in \mathbb{R}$ is a constant, we write $X_n \xrightarrow{p} c$ to denote that X_n converges in probability to c that is, for any $\epsilon > 0$, we have $\mathbb{P}(|X_n - c| > \epsilon) \to 0$ as $n \to \infty$. We write $X_n \xrightarrow{d} X$ to denote that X_n converges in distribution to X. Let $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ be two sequences of real numbers going to infinity as $n \to \infty$. We write $X_n = o_p(a_n)$, if $|X_n|/a_n \xrightarrow{p} 0$, and write $X_n = O_p(a_n)$ if $\mathbb{P}(|X_n| \leqslant C|a_n|) \to 1$ as $n \to \infty$ for some constant C. We write $a_n = o(b_n)$, if $a_n/b_n \to 0$, and write $a_n = O(b_n)$ if for some constant C, $|a_n| \leqslant C|b_n|$. If E_n is a measurable subset of Ω, for any $n \in \mathbb{N}$, we say that the sequence $\{E_n\}_{n\in\mathbb{N}}$ occurs with high probability (w.h.p.) if $\mathbb{P}(E_n) = 1 - o(1)$, as $n \to \infty$. The notation $\mathbb{1}\{E\}$ is used for the indicator of an event E; this is 1 if E holds and 0 otherwise.

Contributions of Chapter 2 : Limit Theorems for Default Contagion and Systemic Risk

We classify the financial institutions into a type set $\mathcal{X}$, which is countable. Without loss of generality, we assume that the institutions belonging to the same type have the same out degree, the same in degree and the same default threshold distribution.

Assumptions 2.1 and 2.2

- *There exists a classification of the financial institutions into a countable set of possible characteristics $\mathcal{X}$ such that, for each $n \in \mathbb{N}$, the institutions in the same characteristic class have the same threshold distribution function (denoted by $q_x^{(n)}$ for institutions in class $x \in \mathcal{X}$).*

- *For a size n network, let $\mu_x^{(n)}$ be the type distribution and $q_x^{(n)}$ be the default threshold distribution of type $x \in \mathcal{X}$. For some probability distribution functions μ and q over the set of characteristics $\mathcal{X}$ and independent of n, we have $\mu_x^{(n)} \to \mu_x$ and $q_x^{(n)}(\theta) \to q_x(\theta)$ as $n \to \infty$, for all $x \in \mathcal{X}$ and $\theta = 0, 1, \ldots, d_x^+$. Moreover, we assume that $\sum_{\theta=0}^{d_x^+} q_x(\theta) = 1$ for all $x \in \mathcal{X}$.*

In order to study the CLT results, we need to restrict our attention to the sparse networks regime.

Assumption 2.3a-b

- *(a) We assume that, as $n \to \infty$, the average degrees converges and is finite:*

$$\lambda^{(n)} := \sum_{x\in\mathcal{X}} d_x^+ \mu_x^{(n)} = \sum_{x\in\mathcal{X}} d_x^- \mu_x^{(n)} \longrightarrow \lambda := \sum_{x\in\mathcal{X}} d_x^+ \mu_x \in (0, \infty).$$

- *(b)We assume that for every constant $A > 1$, we have*

$$\sum_{i=1}^{n} A^{d_i^+} = n \sum_{x\in\mathcal{X}} \mu_x^{(n)} A^{d_x^+} = O(n) \quad \textit{and} \quad \sum_{i=1}^{n} A^{d_i^-} = n \sum_{x\in\mathcal{X}} \mu_x^{(n)} A^{d_x^-} = O(n).$$

Notice that Assumption 2.3b implies 2.3a. The LLN results don't require Assumption 2.3b. We only highlight the CLT results in this introduction.

For $z \in [0,1]$, we define the functions:

$$b(d,z,\ell) := \mathbb{P}(\mathsf{Bin}(d,z) = \ell) = \binom{d}{\ell} z^\ell (1-z)^{d-\ell},$$
$$\beta(d,z,\ell) := \mathbb{P}(\mathsf{Bin}(d,z) \geqslant \ell) = \sum_{r=\ell}^{d} \binom{d}{r} z^r (1-z)^{d-r},$$

where $\mathsf{Bin}(d,z)$ denotes the binomial distribution with parameters d and z. We define further

$$f_S^{(n)}(z) := \sum_{x\in\mathcal{X}} \mu_x^{(n)} \sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big), \quad f_D^{(n)}(z) = 1 - f_S^{(n)}(z),$$
$$f_W^{(n)}(z) := \lambda^{(n)} z - \sum_{x\in\mathcal{X}} \mu_x^{(n)} d_x^- \sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big),$$
$$s_{x,\theta,\ell}^{(n)}(z) := \mu_x^{(n)} q_x^{(n)}(\theta) b\big(d_x^+, 1-z, \ell\big).$$

Unless stated, we define the functions without upscript (n) by replacing the threshold distribution $q_x^{(n)}(\theta)$ and degree distribution $\mu_x^{(n)}$ by their limiting distribution $q_x(\theta)$ and μ_x respectively. For example, define

$$s_{x,\theta,\ell}(z) := \mu_x q_x(\theta) b\big(d_x^+, 1-z, \ell\big).$$

Denote by $D_n(t)$ and $S_n(t)$ the number of defaulted and solvent institutions at time t for size n network, respectively. For $x \in \mathcal{X}, \theta \in \mathbb{N}, \ell = 0, \ldots, \theta - 1$. We let $S_{x,\theta,\ell}^{(n)}(t)$ denote the number of solvent institutions with type x, threshold θ and ℓ defaulted neighbors at time t. Let $W_n(t)$ be the number of infected out half-edges and $\tau_n^\star$ be the stopping time that the default contagion stops. Notice that when there is no more infected out half-edges in the network, the contagion stops.

Let

$$z^\star := \sup\Big\{z \in [0,1] : \lambda z - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=1}^{d_x^+} q_x(\theta)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big) = 0\Big\},$$

which can be viewed as the limit of stopping time $\tau_n^\star$. Let $\widehat{z}_n$ be the largest $z \in [0,1]$ such that $f_W^{(n)}(z) = 0$.

The asymptotic normality at the final time of contagion is given in the following.

Theorem 2.7 *Suppose that Assumptions 2.1-2.3b hold. Let $t^\star = -\ln z^\star$,If $z^\star \in (0,1]$ and $z^\star$ is a stable solution, i.e. $\alpha := f_W'(z^\star) > 0$, then we have jointly*

$$n^{-1/2}(D_n(\tau_n^\star) - n f_D^{(n)}(\widehat{z}_n)) \xrightarrow{d} \mathcal{Z}_D(t^\star) - \alpha^{-1} f_D'(z^\star)\mathcal{Z}_W(t^\star),$$

$$n^{-1/2}(S_n(\tau_n^\star) - n f_S^{(n)}(\widehat{z}_n)) \xrightarrow{d} \mathcal{Z}_S(t^\star) - \alpha^{-1} f_S'(z^\star)\mathcal{Z}_W(t^\star),$$

furthermore, $\widehat{z}_n \xrightarrow{p} z^\star$ *and, for all* $x \in \mathcal{X}$, $0 \leqslant \ell < \theta \leqslant d_x^+$,

$$n^{-1/2}(S^{(n)}_{x,\theta,\ell}(\tau_n^\star) - ns^{(n)}_{x,\theta,\ell}(\widehat{z}_n)) \xrightarrow{d} \mathcal{Z}^*_{x,\theta,\ell}(t^\star) - \alpha^{-1}s'_{x,\theta,\ell}(z^\star)\mathcal{Z}_W(t^\star),$$

where $\mathcal{Z}_D, \mathcal{Z}_S, \mathcal{Z}_W$ *are some centred Gaussian processes with covariance functions characterized explicitly.*

We also study the CLT regarding the systemic risk of the financial network.

System-wide wealth : Let $\bar{\Gamma}^\Diamond_n$ denote the total wealth in the financial system if there is no default in the system. We define the system-wide aggregation function as

$$\Gamma^\Diamond_n(t) := \bar{\Gamma}^\Diamond_n - \sum_{x\in\mathcal{X}} \bar{L}^\odot_x D^{(n)}_x(t) - \sum_{x\in\mathcal{X}} \bar{L}^\Diamond_x \sum_{\theta=1}^{d_x^+}\sum_{\ell=1}^{\theta-1} \ell S^{(n)}_{x,\theta,\ell}(t),$$

where $D^{(n)}_x(t)$ is the number of defaulted institutions of type x. For each type $x \in \mathcal{X}$, we consider a bounded fixed (type-dependent) societal cost $\bar{L}^\odot_x$ for defaulted institutions and a bounded fixed (host institutions' type-dependent) cost $\bar{L}^\Diamond_x$ over each defaulted links.

Assume that $\bar{\Gamma}^\Diamond_n/n \to \bar{\Gamma}^\Diamond$ when the size of network $n \to \infty$. Let us define

$$f^{(n)}_\Diamond(z) := \bar{\Gamma}^\Diamond_n/n - \sum_{x\in\mathcal{X}} \bar{L}^\odot_x f^{(n)}_D(z) - \sum_{x\in\mathcal{X}} \bar{L}^\Diamond_x \sum_{\theta=1}^{d_x^+}\sum_{\ell=1}^{\theta-1} \ell s^{(n)}_{x,\theta,\ell}(z),$$

$$f_\Diamond(z) := \bar{\Gamma}^\Diamond - \sum_{x\in\mathcal{X}} \bar{L}^\odot_x f_D(z) - \sum_{x\in\mathcal{X}} \bar{L}^\Diamond_x \sum_{\theta=1}^{d_x^+}\sum_{\ell=1}^{\theta-1} \ell s_{x,\theta,\ell}(z).$$

The result is as follows.

Theorem 2.8. *Suppose that Assumptions 2.1-2.3a hold. The final (system-wide) aggregation functions satisfy:*

(i) If $z^\star = 0$ *then asymptotically almost all institutions default during the cascade and*

$$\frac{\Gamma^\Diamond_n(\tau_n^\star)}{n} \xrightarrow{p} \bar{\Gamma}^\Diamond - \sum_{x\in\mathcal{X}} \mu_x \bar{L}^\odot_x.$$

(ii) If $z^\star \in (0,1]$ *and* $z^\star$ *is a stable solution, i.e.* $f'_W(z^\star) > 0$, *then* $\frac{\Gamma^\Diamond_n(\tau_n^\star)}{n} \xrightarrow{p} f_\Diamond(z^\star)$ *and, under Assumption 2.3b,*

$$n^{-1/2}\big(\Gamma^\Diamond_n(\tau_n^\star) - nf^{(n)}_\Diamond(\widehat{z}_n)\big) \xrightarrow{d} \mathcal{Z}^\star_\Diamond,$$

where $\mathcal{Z}^\star_\Diamond$ *is a centred Gaussian random variable with variance* $\sigma^\star_\Diamond$ *given by (2.62).*

1.2.3 Fire sales and equilibrium price

As mentioned earlier, when a defaulted institution affects its outgoing neighbors, it may create a currency shortage for those neighbors. Consequently, when an institution is exposed to a defaulted one, it may be compelled to liquidate a certain amount of illiquid assets to maintain its cash reserves according to regulator constraints. Different types of institutions employ different strategies for liquidation. In the context of a financial crisis, fire sales occur when an institution attempts or is forced to sell a substantial quantity of assets within a short period of time.

Chapter 3 focuses on investigating the combined impact of fire sales and default cascades on systemic risk in complex financial networks during a financial crisis. We study the instantaneous fire sales in financial networks, using the results from Chapter 2. The term "instantaneous fire sales" refers to a network that reacts swiftly to an external shock. As soon as the shock occurs, the default cascade and the associated fire sales process are simultaneously triggered. Unlike in [106], where the fire sale prices change round by round, in our study, the fire sale prices are determined at the onset of the shock. Institutions are compelled to liquidate random amounts of illiquid assets to offset interbank losses during the default cascade. The contagion process now depends on the liquidation prices since the default threshold is influenced by the financial characteristics of the institution, and the value of illiquid assets is a component of those characteristics. The financial system aims to achieve a balanced state following the occurrence of the default contagion and fire sales.

The key point is to find an equilibrium price $p_n^\star$ after shock. We consider a conservative approach and assume that the illiquid assets can be sold only at certain final price. After all sales, the market gives a price for illiquid assets according to the inverse demand function g. Let $\Gamma_n(\tau_n^\star(p); p)$ be the total sold amount in the end. Then under the sold price p, the price given by inverse demand function is $g(\Gamma_n(\tau_n^\star(p); p)/n)$, where $\tau_n^\star(p)$ is the final stopping time which depends also on p. This motivates us to define the equilibrium price of the illiquid asset as

$$p_n^\star = \sup\big\{p \in [p_{\min}, p_0] : p \leqslant g(\Gamma_n(\tau_n^\star(p); p)/n)\big\},$$

where p_0 is the initial price without fire sales and $p_{\min}$ is the minimum price with all assets sold. For each fixed sold price $p \in [p_{\min}, p_0]$, we obtain the limit results, both LLN and CLT, regarding the features of networks in the combined process of default contagion and fire sales, such as the sold shares process $\Gamma_n(t; p)$. Under some conditions, we also obtain the limit results (both LLN and CLT) for the equilibrium price $p_n^\star$ and for the features regarding the network structure in the equilibrium state. Finally, we extend the tractable framework to multi-type illiquid assets framework.

Contribution of Chapter 3: Fire Sales and Default Cascades

We first put an assumption on the inverse demand function g.

Assumption 3.1 *Let $p_{\min} \geqslant 0$. We assume that $g : [0, \gamma_{\max}] \to [p_{\min}, p_0]$ satisfies:*

(i) $g(0) = p_0$ (in absence of liquidations the price is given exogenously by p_0).

(ii) $g \in \mathcal{C}^1$ *and it is a non-increasing function of* $x \in [0, \gamma_{\max}]$ *(the price is non-increasing with the average excess supply* x*).*

(iii) $g(\gamma_{\max}) = p_{\min} \geqslant 0$.

Assumption 3.2 *There exists a classification of the financial institutions into a countable set of possible classes* $\mathcal{X}$ *such that, for each* $n \in \mathbb{N}$ *and for all* $p \in [p_{\min}, p_0]$*, the institutions in the same class have the same threshold distribution function (denoted by* $q_x^{(n)}$ *for the institutions in class* $x \in \mathcal{X}$*). Namely, for all* $i \in [n]$ *and all* $\theta \in \mathbb{N}$,

$$\mathbb{P}(\Theta_i^{(n)}(p) = \theta) = q^{(n)}_{x_i^{(n)}}(\theta; p).$$

Assumption 3.3 *For some probability distribution functions* μ *and* $q(.;p)$ *over the set of classes* $\mathcal{X}$ *(independent of* n*), we have* $\mu_x^{(n)} \to \mu_x$ *and* $q_x^{(n)}(\theta;p) \to q_x(\theta;p)$ *as* $n \to \infty$*, for all* $x \in \mathcal{X}, \theta = 0, 1, \dots, d_x^+$ *and* $p \in [p_{\min}, p_0]$*. The empirical threshold distributions satisfy* $q_x^{(n)}(\theta;p) \in \mathcal{C}^1$ *and* $q_x(\theta;p) \in \mathcal{C}^1$ *on* $p \in [p_{\min}, p_0]$*. Moreover, as* $n \to \infty$, $\frac{\partial q_x^{(n)}}{\partial p}(\theta;p)$ *converges uniformly to* $\frac{\partial q_x}{\partial p}(\theta;p)$ *as a function of* p *for all* $x \in \mathcal{X}$ *and* $\theta = 0, 1, \dots, d_x^+$.

In Chapter 3, we also consider the possibility that an institution never defaults, i.e., it remains solvent even if all its counterparties default. We denote such threshold by ∞. We assume each institution liquidates a random amount of illiquid asset as long as it has a defaulted in neighbor. We assume these liquidation are i.i.d. and dependent on type x and threshold θ. Let $\bar{\gamma}_x$ be the constant value of liquidation for each initially defaulted institution with type x.

Assumption 3.4-Liquidation *The mean* $\bar{\ell}_{x,\theta}(p)$ *and variance* $\varsigma^2_{x,\theta}(p)$ *of sold shares for each liquidation are both continuous in* p*, for all* $x \in \mathcal{X}$ *and* $\theta \in \{0, 1, \dots, d_x^+\} \cup \{\infty\}$.

All assumptions above are supposed to hold in Chapter 3. We define the following functions, which are the limit functions of liquidations,

$$f^{(n)}_{x,\theta}(z;p) := \mu_x^{(n)} q_x^{(n)}(\theta;p)\big(\theta - \sum_{\ell = d_x^+ - \theta + 1}^{d_x^+} \beta(d_x^+, z, \ell)\big), \quad f^{(n)}_{x,\infty}(z;p) := (1-z)\mu_x^{(n)} q_x^{(n)}(\infty;p) d_x^+,$$

and,

$$f^{(n)}_{\Gamma}(z;p) := \sum_{x \in \mathcal{X}} \Big(\mu_x^{(n)} \bar{\gamma}_x q_x^{(n)}(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p) f^{(n)}_{x,\theta}(z;p) + \bar{\ell}_{x,\infty}(p) f^{(n)}_{x,\infty}(z;p) \Big).$$

The time-transformed versions for the above functions are then defined as

$$\widehat{f}^{(n)}_{x,\theta}(t;p) := f^{(n)}_{x,\theta}(e^{-t};p), \qquad \widehat{f}^{(n)}_{\Gamma}(t;p) := f^{(n)}_{\Gamma}(e^{-t};p),$$

and the same for other functions.

Let $\Gamma_n(t;p)$ be the total shares of illiquid assets sold by time t.

We define

$$z_n^\star(p) := \sup\big\{z \in [0,1] : f_W^{(n)}(z;p) = 0\big\},$$

and

$$z^\star(p) := \sup\big\{z \in [0,1] : f_W(z;p) = 0\big\},$$

where $f_W^{(n)}(z;p)$ and $f_W(z;p)$ are defined the same as in previous section but now it is price dependent. We then let $t^\star(p) := -\ln z^\star(p)$ and $t_n^\star(p) := -\ln z_n^\star(p)$.

Let $f^1(z;p)$ and $f^2(z;p)$ denote the partial derivative w.r.t. the first parameter and second parameter respectively. We have the following theorem regarding the asymptotic normality of the final total sold shares.

Theorem 3.15 *For any fixed $p \in [p_{\min}, p_0]$, as $n \to \infty$, the final total sold shares satisfy:*

(i) Under Assumption 2.3a, if $z^\star(p) = 0$, then asymptotically almost all institutions default after shock and (as $n \to \infty$)

$$\frac{\Gamma_n(\tau_n^\star;p)}{n} \xrightarrow{p} \sum_{x\in\mathcal{X}} \mu_x\Big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta;p)\Big).$$

(ii) Under Assumption 2.3b, if $z^\star(p) \in (0,1]$ and $z^\star(p)$ is a stable solution, i.e., $\alpha(p) := f_W^1(z^\star(p);p) > 0$, then

$$n^{-1/2}(\Gamma_n(\tau_n^\star;p) - n\widehat{f}_\Gamma^{(n)}(t_n^\star(p);p)) \xrightarrow{d} \mathcal{Z}_\Gamma(t^\star(p);p) - \alpha(p)^{-1} f_\Gamma^1(z^\star(p);p)\mathcal{Z}_W(t^\star(p);p),$$

where $\mathcal{Z}_\Gamma(t;p)$ and $\mathcal{Z}_W(t;p)$ are some Gaussian processes depending on p.

We also show the following limit theorem on the price given by inverse demand function $\kappa_n(p) := g(\Gamma_n(\tau_n^\star(p);p)/n)$.

Theorem 3.17 *For any $p \in [p_{\min}, p_0]$ fixed and as $n \to \infty$, the price $\kappa_n(p)$ given by the inverse demand function satisfies:*

(i) Under Assumption 2.3a, if $z^\star(p) = 0$ then asymptotically almost all institutions default after shock and

$$\kappa_n(p) \xrightarrow{p} g\Big(\sum_{x\in\mathcal{X}} \mu_x\big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta;p)\big)\Big).$$

(ii) Under Assumption 2.3b, if $z^\star(p) \in (0,1]$ and $z^\star(p)$ is a stable solution, i.e., $\alpha(p) := f_W^1(z^\star(p);p) > 0$, then

$$n^{1/2}\big(\kappa_n(p) - g\big(\widehat{f}_\Gamma^{(n)}(t_n^\star(p);p)\big)\big) \xrightarrow{d} g'\big(f_\Gamma(z^\star(p);p)\big)\Big[\mathcal{Z}_\Gamma(t^\star(p);p) - \alpha(p)^{-1} f_\Gamma^1(z^\star(p);p)\mathcal{Z}_W(t^\star(p);p)\Big],$$

where g' denotes the first derivative of g.

We next obtain limit theorem for the equilibrium price after shock. For the network of size n, we define

$$\bar{p}_n := \sup\big\{p \in [p_{\min}, p_0] : p \leqslant g\big(f_\Gamma^{(n)}(z_n^\star(p); p)\big)\big\}.$$

Similarly define its limit counterpart

$$\bar{p} := \sup\big\{p \in [p_{\min}, p_0] : p \leqslant g\big(f_\Gamma(z^\star(p); p)\big)\big\}. \tag{1.1}$$

We say that $\bar{p}$ is a *stable* fixed point solution if either $\bar{p} = p_{\min}$ or, $\bar{p} \in (p_{\min}, p_0]$ and there exists some $\epsilon > 0$ such that $p < g\big(f_\Gamma(z^\star(p); p)\big)$ for all $p \in (\bar{p} - \epsilon, \bar{p})$.

The result regarding the equilibrium price is the following.

Theorem 3.18 *As $n \to \infty$, the equilibrium price satisfies:*

(i) Under Assumption 2.3a, if $z^\star(\bar{p}) = 0$ and $\bar{p}$ is a stable solution, then the equilibrium price converges to $p_n^\star \xrightarrow{p} \bar{p}$, where $\bar{p}$ is the largest solution of the fixed point equation

$$p = g\Big(\sum_{x\in\mathcal{X}} \mu_x\big(\bar{\gamma}_x q_x(0; p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta; p)\big)\Big).$$

(ii) Under Assumption 2.3b, if $z^\star(\bar{p}) \in (0, 1]$ is a stable solution of $f_W(z; \bar{p}) = 0$, i.e., $\alpha(\bar{p}) := f_W^1(z^\star; \bar{p}) > 0$, and $\bar{p}$ is a stable solution of (1.1), then

$$n^{1/2}(p_n^\star - \bar{p}_n) \xrightarrow{d} -\rho^{-1}(\bar{p})\mathcal{Z}_V(\bar{p}),$$

where

$$\rho(p) := 1 - g'\big(f_\Gamma(z^\star(p); p)\big)\Big[-f_\Gamma^1(z^\star(p); p)\alpha(p)^{-1} f_W^2(z^\star(p); p) + f_\Gamma^2(z^\star(p); p)\Big],$$

and,

$$\mathcal{Z}_V(p) := -g'\big(f_\Gamma(z^\star; p)\big)\Big[\mathcal{Z}_\Gamma(t^\star(p); p) - \alpha(p)^{-1} f_\Gamma^1(z^\star; p)\mathcal{Z}_W(t^\star(p); p)\Big]$$

is a Gaussian random variable with mean 0.

1.3 Risk processes on networks

Chapter 4 is dedicated to the study of the risk processes on networks. Let recall some background on classical risk processes.

1.3.1 Classical risk processes

The classical compound risk process with Poisson claim arrivals, or the Cramér-Lundberg model ([100, 171]) has been extensively used in quantitative risk management, see e.g., [103, 173]. In this model, the aggregate capital of an insurer who starts with initial capital γ, premium rate α and (loss) claim sizes (L_k) is given by the following spectrally negative compound Poisson process

$$C(t) = \gamma + \alpha t - \sum_{k=1}^{\mathcal{N}(t)} L_k,$$

where $L_k, k \in \mathbb{N}$, are i.i.d. non-negative random variable following a distribution F with mean $\bar{L}$ and $\mathcal{N}(t)$ is a Poisson process with intensity $\beta > 0$ independent of L_k. The ruin time for the insurer with initial capital γ is defined by

$$\tau(\gamma) := \inf\{t \mid C(t) \leqslant 0\},$$

(with the convention that $\inf \varnothing = \infty$) and the central question is to find the ruin probability

$$\psi(\gamma) := \mathbb{P}(\tau(\gamma) < \infty).$$

It is known (see e.g. [35, 115]) that whenever $\beta\bar{L} > \alpha$, we have $\psi(\gamma) = 1$ for all $\gamma \in \mathbb{R}$ and whenever $\beta\bar{L} < \alpha$, the ruin probability can be computed using the famous Pollaczek–Khinchine formula as

$$\psi(\gamma) = \left(1 - \frac{\beta\bar{L}}{\alpha}\right) \sum_{k=0}^{\infty} \left(\frac{\beta\bar{L}}{\alpha}\right)^k \left(1 - \widehat{F}^{*k}(\gamma)\right),$$

where

$$\widehat{F}(\gamma) = \frac{1}{\bar{L}} \int_0^{\gamma} \big(1 - F(u)\big) du,$$

and the operator $(\cdot)^{*k}$ denotes the k-fold convolution.

1.3.2 Risk processes on financial networks

Recent efforts have been dedicated to the study of risk processes on networks. In [56], the authors consider risk processes and ruin probabilities in bipartite networks. But it is more like a linear combination of several classical risk process with certain independence. In Chapter 4, we study a more general risk model on heterogeneous financial networks, where institutions can receive capital recovery in time, i.e. there exists a non-decreasing function of time $\alpha_i(t)$ for each agent $i \in [n]$ in the network. Let $C_i(t)$ be the total capital of agent i at time t. We consider the stochastic networked risk process as following, for each agent $i \in [n]$:

$$C_i(t) := \gamma_i(1 - \epsilon_i) + \alpha_i(t) - \delta_i - \sum_{j \in [n]: j \to i} L_{ji} \mathbb{1}\{\tau_j + T_{ji} \leqslant t\}, \tag{1.2}$$

where $\tau_j := \inf\{t : C_j(t) \leqslant 0\}$ denotes the ruin time for agent $j \in [n]$ and L_{ji} is the random interbank loss brought by j when it defaults, γ_i is amount of external assets exposed to risk, ϵ_i is the shock (lost fraction of external assets), T_{ji} is the time delay that the loss L_{ji} happens to i and δ_i represents the total value of claims held by end-users on agent i (deposits). In Chapter 2, we study the default contagion without capital recovery, where the default threshold only depends on the capital profile and received interbank losses. In this situation, the default threshold of each institution has a fixed distribution. In Chapter 4, the threshold varies in time with capital recovery, and the analysis becomes more sophisticated. In [28], the authors study the contagion with recovery in a similar setting, but with the recovery on the threshold and being of a special form. Here we study a more general case.

We investigate the ruin probability for risk processes on large-scale networks. We establish LLN results for network structures by using a probabilistic approach, which relies on knowledges on the Glivenko-Cantelli class and associated theorems. Our study encompasses various aspects of networked risk processes. Specifically, we study the limit theorems related to the contagion dynamics and the networked ruin probabilities for risk processes within a stochastic network setting. We also provide estimations for ruin probabilities for complex networked risk processes, which involve both losses coming from network and heterogeneous losses originating from external sources.

Contribution of Chapter 4: Ruin probabilities for Risk Processes in Stochastic Networks

We consider a general loss reveal intensity process, denoted by $\mathcal{R}_n(t)$, to describe the intensity of internetwork loss reveals. Specifically, if a loss is revealed at time $t_1 \in \mathbb{R}_+$, we wait for an exponential time with parameter $\mathcal{R}_n(t_1)$ until the next loss reveal.

Assumption 4.1 *We assume that for some probability distribution μ over $\mathcal{X}$ and independent of n, we have that $\mu_x^{(n)} \to \mu_x$, as $n \to \infty$, for all $x \in \mathcal{X}$.*

Assumption 4.3 *We assume that the loss intensity function $\mathcal{R}_n$ satisfies $\mathcal{R}_n(t) = 0$ for $t > \tau_n^\star$, and $\mathcal{R}_n(t) = n\mathfrak{R}(t) + o_p(n)$ for $t \leqslant \tau_n^\star$ with $\mathfrak{R}(t)$ continuous, positive and $o_p(n)$ is uniform for $t \leqslant \tau_n^\star$.*

For each $x \in \mathcal{X}$, we denote by $\boldsymbol{L}_x := (L_x^{(1)}, \dots, L_x^{(d_x^+)})$ the sequence of independent random losses with distribution F_x and let $\boldsymbol{\ell}_x = (\ell_x^{(1)}, \ell_x^{(2)}, \dots, \ell_x^{(d_x^+)})$ be a realization of $\boldsymbol{L}_x$. For a given $\boldsymbol{\ell}_x$ and a given initial shock ϵ_x, $\tau_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x)$ is defined as the time threshold for default, namely if the loss happens before this time threshold, then the agent defaults. It is formally defined as

$$\tau_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x) := \inf\Big\{t \geqslant 0 : \gamma_x(1-\epsilon_x) + \alpha_x(t) - \delta_x \geqslant \sum_{i=1}^{\theta} \ell_x^{(i)}\Big\}.$$

For a given positive density function $\mathfrak{R} : \mathbb{R}_0^+ \to \mathbb{R}^+$ with $\|\mathfrak{R}\|_{L^1} < \infty$, $x \in \mathcal{X}$ and $\theta = 0, 1, \dots, d_x^+$, we let the survival probability be (for all $t \geqslant 0$)

$$P_{x,\theta}^{\mathfrak{R}}(t, \epsilon_x, \boldsymbol{\ell}_x) := \mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{\ell}_x) = 0, U_{(1)}^{\mathfrak{R},t} > \tau_{x,1}(\epsilon_x, \boldsymbol{\ell}_x), \dots, U_{(\theta)}^{\mathfrak{R},t} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x)),$$

with the convention $P_{x,0}^{\mathfrak{R}}(t, \epsilon_x, \boldsymbol{\ell}_x) := \mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{\ell}_x) = 0)$ for all $x \in \mathcal{X}$, where $U_{(1)}^{\mathfrak{R},t}, U_{(2)}^{\mathfrak{R},t}, \dots, U_{(\theta)}^{\mathfrak{R},t}$ are the order statistics of θ i.i.d. random variables $\{U_i^{\mathfrak{R},t}\}_{i=1,\dots,\theta}$ with distribution

$$\mathbb{P}(U_i^{\mathfrak{R},t} \leqslant y) = \frac{\int_0^y \mathfrak{R}(s)ds}{\int_0^t \mathfrak{R}(s)ds}, \qquad y \leqslant t.$$

The survival probability at time t for any agent of type x with θ incoming losses absorbed by t is denoted by $\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)$, and is defined as

$$\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t) := \mathbb{E}[P^{\mathfrak{R}}_{x,\theta}(t, \epsilon_x, \boldsymbol{L}_x)] = \mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{L}_x) = 0, U^{\mathfrak{R},t}_{(1)} > \tau_{x,1}(\epsilon_x, \boldsymbol{L}_x), \ldots, U^{\mathfrak{R},t}_{(\theta)} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{L}_x)),$$

Let $\mathbb{R}^+_0$ be the half line $[0, \infty)$. For a given positive function $\mathfrak{R} : \mathbb{R}^+_0 \to \mathbb{R}^+$, we define

$$\phi^{\mathfrak{R}}(t) := \frac{\int_0^{t \wedge t_{\mathfrak{R}}(\lambda)} \mathfrak{R}(s) ds}{\lambda},$$

where

$$t_{\mathfrak{R}}(\lambda) := \inf\{t \geqslant 0 : \int_0^t \mathfrak{R}(s) ds \geqslant \lambda\},$$

if $\|\mathfrak{R}\|_{L^1} \leqslant \lambda$ we set $t_{\mathfrak{R}}(\lambda) := \infty$. We also define

$$f^{\mathfrak{R}}_S(t) := \sum_{x \in \mathcal{X}} \mu_x \sum_{\theta=0}^{d^+_x} b(d^+_x, \phi^{\mathfrak{R}}(t), \theta) \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t), \quad f^{\mathfrak{R}}_D(t) = 1 - f^{\mathfrak{R}}_S(t),$$

and

$$f^{\mathfrak{R}}_W(t) := \lambda(1 - \phi^{\mathfrak{R}}(t)) - \sum_{x \in \mathcal{X}} \mu_x d^-_x \sum_{\theta=0}^{d^+_x} b(d^+_x, \phi^{\mathfrak{R}}(t), \theta) \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t).$$

The main result is the following.

Theorem 4.5 *Under Assumption 2.3a and 4.1, for any given loss intensity function $\mathcal{R}_n$ satisfying Assumption 4.3, we have as $n \to \infty$,*

$$\sup_{t \leqslant \tau^\star_n} \Big| \frac{S^{(n)}_{x,\theta}(t)}{n} - \mu_x b(d^+_x, \phi^{\mathfrak{R}}(t), \theta) \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t) \Big| \xrightarrow{p} 0.$$

Further, as $n \to \infty$,

$$\sup_{t \leqslant \tau^\star_n} \Big| \frac{S^{(n)}(t)}{n} - f^{\mathfrak{R}}_S(t) \Big| \xrightarrow{p} 0, \quad \sup_{t \leqslant \tau^\star_n} \Big| \frac{D^{(n)}(t)}{n} - f^{\mathfrak{R}}_D(t) \Big| \xrightarrow{p} 0,$$

and the process W_n satisfies

$$\sup_{t \leqslant \tau^\star_n} \Big| \frac{W_n(t)}{n} - f^{\mathfrak{R}}_W(t) \Big| \xrightarrow{p} 0.$$

Define

$$t^\star_{\mathfrak{R}} := \inf\big\{t \in [0, 1] : f^{\mathfrak{R}}_W(t) = 0\big\}.$$

We say that $t^\star_{\mathfrak{R}} < \infty$ is a stable solution of $f^{\mathfrak{R}}_W(t) = 0$ if there exists a small $\epsilon > 0$ such that $f^{\mathfrak{R}}_W(t)$ is negative on $[t^\star_{\mathfrak{R}}, t^\star_{\mathfrak{R}} + \epsilon)$.

The result for final defaults is the following. Recall that $\tau^\star_n$ is the stopping time that the loss propagation process comes to the end.

Theorem 4.9 *Under Assumption 2.3a and 4.1, and for any given loss intensity function $\mathcal{R}_n$ satisfying Assumption 4.3, we have as $n \to \infty$:*

(i) If $\int_0^{t^\star_{\mathfrak{R}}} \mathfrak{R}(s)ds = \lambda$, then asymptotically all agents are ruined by the end of the loss propagation process, i.e.

$$D^{(n)}(\tau^\star_n) = n - o_p(n).$$

(ii) If $t^\star_{\mathfrak{R}} < \infty$ is a stable solution of $f^{\mathfrak{R}}_W(t) = 0$ and $\int_0^{t^\star_{\mathfrak{R}}} \mathfrak{R}(s)ds < \lambda$, then the ruin probability of an agent of type $x \in \mathcal{X}$ converges to

$$\frac{D^{(n)}_x(\tau^\star_n)}{n\mu^{(n)}_x} \xrightarrow{p} 1 - \sum_{\theta=0}^{d^+_x} b(d^+_x, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t^\star_{\mathfrak{R}}),$$

and the total number of ruined agents satisfies

$$D^{(n)}(\tau^\star_n) = n \sum_{x \in \mathcal{X}} \mu_x (1 - \sum_{\theta=0}^{d^+_x} \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t^\star_{\mathfrak{R}})) + o_p(n).$$

(iii) If $t^\star_{\mathfrak{R}} = \infty$ and $\|\mathfrak{R}\|_{L^1} < \lambda$, then the ruin probability of an agent of type $x \in \mathcal{X}$ converges to

$$\frac{D^{(n)}_x(\tau^\star_n)}{n\mu^{(n)}_x} \xrightarrow{p} 1 - \sum_{\theta=0}^{d^+_x} b\big(d^+_x, \|\mathfrak{R}\|_{L^1}/\lambda, \theta\big)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(\infty),$$

and the total number of ruined agents satisfies

$$D^{(n)}(\tau^\star_n) = n \sum_{x \in \mathcal{X}} \mu_x (1 - \sum_{\theta=0}^{d^+_x} b\big(d^+_x, \|\mathfrak{R}\|_{L^1}/\lambda, \theta\big)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(\infty)) + o_p(n),$$

where $\mathcal{S}^{\mathfrak{R}}_{x,\theta}(\infty)$ denotes the limit of $\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)$ as $t \to \infty$.

Assumption 4.4 *We assume that, as $n \to \infty$, $\sum_{i \in [n]} (d^+_i + d^-_i)^2 = O(n)$.*

The aforementioned results were derived under the assumption that the loss reveal intensity function is known. In the following result, we consider a special case where the loss reveal intensity is dependent on the current number of unrevealed infected outgoing half-edges ($W_n(t)$) within the network.

Theorem 4.10 *Let $\mathbb{L}_\lambda(\mathbb{R}^+)$ be the space of all continuous positive integrable functions f with $\|f\|_1 \leqslant \lambda$. Suppose that the loss reveal intensity satisfies $\mathcal{R}_n(t) = \beta W_n(t)$ for some constant β and the network sequence $\{\mathcal{G}^{(n)}\}_{n \in \mathbb{N}}$ satisfies Assumptions 4.1 and 4.4. Then we have:*

(i) There exists a unique solution $\mathfrak{R}^\star$ *in* $\mathbb{L}_\lambda(\mathbb{R}^+)$ *with an initial value* $\mathfrak{R}^\star(0) = \beta \sum_{x\in\mathcal{X}} \mu_x d_x^-(1-q_{x,0})$ *to the fixed point equation* $\mathfrak{R} = \beta\Psi(\mathfrak{R})$, *where* $\Psi : \mathbb{L}_\lambda(\mathbb{R}^+) \to \mathbb{L}_\lambda(\mathbb{R}^+)$ *is the map*

$$\Psi(\mathfrak{R})(t) = \lambda(1-\phi^{\mathfrak{R}}(t)) - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t).$$

(ii) As $n \to \infty$, *we have*

$$\sup_{t\leqslant\tau_n^\star} \Big|\frac{\beta W_n(t)}{n} - \mathfrak{R}^\star(t)\Big| \xrightarrow{p} 0,$$

and consequently,

$$\sup_{t\leqslant\tau_n^\star} \Big|\frac{S^{(n)}(t)}{n} - f_S^{\mathfrak{R}^\star}(t)\Big| \xrightarrow{p} 0 \quad \text{and} \quad \sup_{t\leqslant\tau_n^\star} \Big|\frac{D^{(n)}(t)}{n} - f_D^{\mathfrak{R}^\star}(t)\Big| \xrightarrow{p} 0.$$

1.4 Graphon mean field games and interacting systems

Systemic risk and mean field systems. Agents in networks are usually influenced by a group of agents in this network, who are the "neighbors" in the specific sense according to the models and contexts. Sometimes, such impact may depend on the whole population. The study of systemic risk and default contagion in complex financial networks has been increasingly connected to the theory of mean field systems and mean field games in recent years, see e.g. [32, 85, 136]. Among them, [85] studies a model of inter-bank borrowing and lending. In [32], the authors study a more complex model of intra-and-inter-bank borrowing and lending, which includes different groups of banks, and the financial impacts come from both the inter-group banks and the banks of different groups. A dynamic mean field model for studying systemic risk and contagion cascade is proposed in [136]. The default cascades can be modelled by an alternative framework. Consider a diffusion dynamic to describe the capital evolution for each agent. Then the default time can be captured by using hitting times, e.g. the hitting time to 0 of the diffusion. One can establish a connection between the proportion of solvent agents in large financial networks and the probability of default in the McKean-Vlasov equation as the size of agents n tends to infinity, see [49, 50, 177]. Overall, mean field particle systems are adapted well to model the evolution of objects of interest in finite networks, and their limit counterparts when $n \to \infty$ can, in turn, give insights into the properties or behaviors of the financial events that we are interested in.

The study of mean-field systems with homogeneous interactions has a rich history, dating back from the works of Boltzmann, Vlasov, McKean and others (see e.g., [33, 154, 172]). They can be viewed as limits of interacting particles systems, originally coming from models in statistical physics. Similar interacting models have been considered for a broad range of applications in different fields, including banking networks, biology, social sciences, etc (see e.g. [74, 135, 136]). Backward Stochastic Differential Equations (BSDEs) of mean-field type have been early studied in [72, 73]. In addition, the theory of mean-field games, introduced by Lasry and Lions in [163] and Huang, Caines and Malhamé [141, 142], has raised significant attention in recent years.

However, a limitation of mean field games is the assumption of homogeneity in interactions, which may not capture the heterogeneity observed in real-world systems. In order to capture the heterogeneity of interaction, systems with multi-type populations have been proposed in many domains, see e.g. [75, 178]. More recently, the study of mean field systems on large networks have been attracted increasing attention, see [52, 61, 97, 110, 155] and the references therein.

In many real-world systems, including financial networks, heterogeneity is prevalent, as different participants have varying levels of influence on each other within the system. This heterogeneity arises from both the underlying graph structure of the system and the diverse characteristics of the players involved. To better model the heterogeneous interactions in such systems, the study on graphon mean field interacting systems has emerged. Graphons, introduced by Lovász in [170], serve as natural continuum limit objects for large and dense graphs, providing a powerful tool for modeling and analyzing complex systems with heterogeneous interactions. The concept of graphon mean field systems has been proposed and increasingly studied in recent years, starting from the well-posedness and large population convergence from particle systems to limit graphon systems (see [47, 60] for forward systems, and [55] for coupled forward-backward systems), to the concentration bounds and concentration of measures of graphon particle systems (see [51, 54]).

Besides the classical interacting diffusions driven by Brownian motions, interacting systems driven by Poisson random measures are also studied in e.g. [3, 52]. In [3], the author investigates multivariate Hawkes processes on heterogeneous graphs and their graphon limits. Incorporating the underlying graph structure into the dynamics is studied in [52]. The use of graphons to analyze heterogeneous interaction in the theory of mean field games is also increasing studied, see [36, 82, 162]. Furthermore, using graphons to learn mean field games on heterogeneous networks has emerged recently, see e.g. [101, 140]. In the second part of this thesis, we first focus on the study of a pure backward graphon mean field system with jumps and its associated risk measures. Then we investigate stochastic control problems based on forward graphon mean field systems with jumps.

Graphons. A graphon is defined as a symmetric measurable function $G : I \times I \to I$, with $I = [0, 1]$. Graphons can be regarded as the limits of edge matrices of weighted graphs, when the size of the graph (number of vertices) goes to infinity. Indeed, by relabelling vertices of the graph by i/n, $i \in [n] := \{1, \dots, n\}$, as n becomes large, the labels $i/n, i \in [n]$ become close to each other, tending to a continuum in $[0, 1]$. Let $\mathcal{B}(I)$ be the Borel algebra on I. The so-called cut norm of a graphon is defined by

$$\|G\|_\square := \sup_{A,B\in\mathcal{B}(I)} \Big| \int_{A\times B} G(u, v)dudv \Big|.$$

We can also view a graphon as an operator from $L^\infty(I)$ to $L^1(I)$, associating any $\phi \in L^\infty(I)$ with:

$$G\phi(u) := \int_I G(u, v)\phi(v)dv.$$

By Lovász [170, Lemma 8.11], the resulting operator norm turns out to be equivalent to the cut norm

$$\|G\|_\square \leqslant \|G\|_{\infty\to1} \leqslant 4\|G\|_\square,$$

with

$$\|G\|_{\infty\to 1} := \sup_{|\phi|\leqslant 1} \|G\phi\|_{L^1}.$$

These norms will be used in the studies of convergence theorems for the graphon systems induced by a sequence of graphons. To study stronger convergence results, we need to consider another operator norm for graphons, regarding G as an operator from $L^\infty(I)$ to $L^\infty(I)$ with the norm defined by

$$\|G\|_{\infty\to\infty} := \sup_{|\phi|\leqslant 1} \|G\phi\|_{L^\infty}.$$

With a given metric space $\mathcal{S}$, denote by $\mathcal{M}_+(\mathcal{S})$ the set of nonnegative Borel measurable measures on $\mathcal{S}$ and denote by $\mathcal{M}^+_{\text{Unif}}([0,1]\times\mathcal{S})$ the set of nonnegative Borel measures on $[0,1]\times\mathcal{S}$ with uniform first marginal. We define the measure-valued function $\Lambda\mu : [0,1] \to \mathcal{M}_+(\mathcal{S})$ for any $\mu \in \mathcal{M}^+_{\text{Unif}}([0,1]\times \mathcal{S})$ as follows:

$$\Lambda\mu(u) := \int_{[0,1]\times\mathcal{S}} G(u,v)\delta_x\mu(dv,dx), \tag{1.3}$$

where δ_x denotes the Dirac measure concentrated at x. For any bounded measurable function $\phi : \mathcal{S} \to \mathbb{R}$, the usual inner product is defined by

$$\langle \Lambda\mu(u), \phi\rangle = \int_{[0,1]\times\mathcal{S}} G(u,v)\phi(x)\mu(dv,dx).$$

Systemic risk on large heterogeneous networks. In Part I, even though we classify the financial institutions through a characteristics set $\mathcal{X}$, the network has equal connection probability between institutions of different types. A significant extension would be to introduce heterogenous connection probability between institutions, which can be modelled by a choice distribution, say $Q_x(\cdot)$ over the set $\mathcal{X}$ for each type $x \in \mathcal{X}$, i.e. for each connection opportunity, a type x institution tends to choose a type y institution with probability $Q_x(y)$ independently. Then the connection probability between a pair of type (x,y) is $Q_x(y)Q_y(x)$, which can be rewritten as some symmetric function $\bar{Q}(x,y)$ over $\mathcal{X}^2$. Here the function $\bar{Q}(x,y)$ plays a similar role as the graphon G, showing the relevance to introduce graphons in the study of systemic risk in large heterogeneous networks. In such framework, the study of percolation in heterogeneous graph is interesting. In [48], the authors study the k-core problem in percolated dense heterogeneous graph sequences converging in the sense of cut norm. This may be related to the study of risk models in large heterogeneous networks. Future work may include some extensions of the models in Part I involving graphons. The study of games between institutions (e.g. optimal connectivity, optimal connection probability) or optimization problems involving an outside regulator (target interventions) are also interesting topics to study. In Part II, we focus on the graphon mean field systems and games. The study of systemic risk in complex graphon mean field models is left for future work.

1.4.1 Graphon mean field backward stochastic differential equations with jumps

Notation and setting. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space. Let $I = [0,1]$ and $\{W_u : u \in I\}$ be a family of i.i.d. one dimensional Brownian motions defined on $(\Omega,\mathcal{F},\mathbb{P})$. Let $\{N_u(dt,de) : u \in$

$I\}$ be a family of independent Poisson measures defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with compensator $\nu_u(de)dt$ such that ν_u is a σ-finite measure on $E := \mathbb{R}_*$, with $\mathbb{R}_* := \mathbb{R}\backslash\{0\}$, equipped with its Borelian σ-algebra $\mathcal{B}(E)$, for each $u \in I$. Let $\{\widetilde{N}_u(dt, de) : u \in I\}$ be their compensator processes. Let $\mathbb{F} = \{\mathcal{F}_t, t \geqslant 0\}$ be the natural filtration associated with $\{W_u : u \in I\}$ and $\{N_u(dt, de) : u \in I\}$. Given a Polish space $\mathcal{S}$, denote by $\mathcal{D}([0,T], \mathcal{S})$ the space of RCLL (right continuous with left limits) functions from $[0,T]$ to $\mathcal{S}$, equipped with the topology of uniform convergence. Let $\mathcal{D} := \mathcal{D}([0,T], \mathbb{R})$. Denote by $\mathcal{P}(\mathcal{S})$ the space of probability measures on $\mathcal{S}$. For a random variable X, $\mathcal{L}(X)$ denotes the law of X. Denote Unif$[0,1]$ the uniform measure on $[0,1]$ and further denote $\mathcal{P}_{\text{Unif}}([0,1] \times \mathcal{S})$ the set of Borel probability measures on $[0,1] \times \mathcal{S}$ with uniform first marginal. We equip all spaces of measure with the topology of weak convergence. Denote by $\mathcal{W}_2$ the Wasserstein-2 distance. For a family of objects $\{X_u\}_{u \in I}$ or a sequence of objects $\{X_i\}_{i \in [n]}$, we use X to represent them for notation simplicity when the context is clear.

In Chapter 5, we study graphon mean field backward stochastic differential equations (BSDEs) with jumps and the associated dynamic risk measures. We consider a backward system. Forward and forward-backward graphon mean field systems without jumps have been studied in [47] and [55] respectively. We consider the following graphon mean-field BSDE with jumps:

$$
\begin{aligned}
X_u(t) = \xi_u &+ \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f(s, x, X_u(s), Z_u(s), \ell_{u,s}(\cdot)) \mu_{y,s}(dx) dy ds - \int_t^T Z_u(s) dW_u(s) \\
&- \int_t^T \int_E \ell_{u,s}(e) \widetilde{N}_u(ds, de), \qquad\qquad u \in I, \quad \text{for} \quad t \in [0,T],
\end{aligned} \tag{1.4}
$$

where $\mu_y := \mathcal{L}(X_y) \in \mathcal{P}(\mathcal{D})$ and $\mu_{y,s} := \mathcal{L}(X_y(s)) \in \mathcal{P}(\mathbb{R})$. We assume that for each $u \in I$, $\xi_u \in L^2(\mathcal{F}_T)$ and the map $u \mapsto \xi_u$ is measurable.

The heterogeneous interaction is governed by the graphon term G. Note that if the interaction is homogeneous, then $G(u,v) \equiv 1$ for all $(u,v) \in [0,1] \times [0,1]$. In this case, the above BSDE with jumps degenerates to the standard mean field case,

$$
\begin{aligned}
X_u(t) = \xi_u &+ \int_t^T \int_I \int_{\mathbb{R}} f(s, x, X_u(s), Z_u(s), \ell_{u,s}(\cdot)) \mu_{y,s}(dx) dy ds - \int_t^T Z_u(s) dW_u(s) \\
&- \int_t^T \int_E \ell_{u,s}(e) \widetilde{N}_u(ds, de), \qquad\qquad u \in I, \quad \text{for} \quad t \in [0,T].
\end{aligned}
$$

This particular case has been studied in [89]. In our research, we conduct a comprehensive analysis of graphon mean field BSDEs with jumps. We establish some fundamental results, including the existence and uniqueness of solutions, estimate for solutions, and comparison theorems. In addition, we also explore the propagation of chaos of its associated N-particle systems. Specifically, we consider an N-coupled BSDE system, where each equation is indexed by $i = 1, \ldots, N$, and has the following form:

$$\begin{aligned}
X_i^N(t) &= \xi_i^N + \int_t^T \frac{1}{N}\sum_{j=1}^N \zeta_{ij}^N f(s, X_j^N(s), X_i^N(s), Z_i^N(s), \ell_s^{N,i}(\cdot))ds - \int_t^T Z_i^N(s)d\widehat{W}_i(s) \\
&\quad - \int_t^T \int_E \ell_s^{N,i}(e)\widetilde{\widehat{N}}_i(ds,de), \quad t \in [0,T] \\
X_i^N(T) &= \xi_i^N,
\end{aligned} \tag{1.5}$$

where $\widehat{W}_i := W_{\frac{i}{N}}$ are i.i.d. Brownian motions, and $\widehat{N}_i(dt,de) = N_{\frac{i}{N}}(dt,de)$ are independent Poisson random measures. We assume that $\xi_i^N \in L^2(\mathcal{F}_T)$ for all $i = 1,\dots,N$. Hereby, $\zeta_{ij}^N : \mathbb{N}\times\mathbb{N} \mapsto \mathbb{R}_0^+$ is symmetric describing the strength of interaction between particle i and j.

The graphon G can be regarded as the limit of ζ_{ij}^N as $N \to \infty$. We study the following two different types of convergence for solution:

- The average type:

$$\frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right].$$

- The maximum type:

$$\max_{i\in[N]} \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right].$$

We introduce the graphon dynamic risk measures induced by the solutions of the graphon mean field BSDEs with jumps. We extend several properties of dynamic risk measures to the graphon mean field case, which include the consistency, continuity, homogeneity, translation invariance, monotonicity, convexity, and absence of arbitrage. These properties have been previously studied in the context of dynamic risk measures without mean field term in [183] and with mean field term in [89]. Additionally, we provide a dual representation formula, a fundamental result in the theory of convex risk measures. This formula provides a methodology for computing dynamic risk measures by taking the supremum over a set of expectations under a family of probability measures. Through an involved proof, we establish the dual representation formula for the graphon dynamic risk measures.

Contribution of Chapter 5 : Graphon Mean Field BSDEs and Associated Dynamic Risk Measures

We introduce the following sets.

- $L^2(\mathcal{F}_t)$ is the set of all $\mathcal{F}_t$-measurable and square integrable random variables, for $t \in [0,T]$.

- $\mathbb{H}^2$ is the set of real-valued predictable processes ϕ such that

$$\|\phi\|_{\mathbb{H}^2} := (\mathbb{E}[\int_0^T \phi_t^2 dt])^{1/2} < \infty.$$

- $L^2_{\nu_u}$ (for each $u \in I$) is the set of all measurable functions $\ell : E \mapsto \mathbb{R}$ such that

$$\|\ell\|_{\nu_u} := (\int_E |\ell(y)|^2 \nu_u(dy))^{1/2} < \infty.$$

  Note that $L^2_{\nu_u}$ is a Hilbert space equipped with the scalar product

$$\langle \ell_1, \ell_2 \rangle_{\nu_u} := \int_E \ell_1(y)\ell_2(y)\nu_u(dy).$$

- $\mathbb{H}^2_{\nu_u}$ (for each $u \in I$) is the set of all predictable processes ℓ such that

$$\|\ell\|_{\mathbb{H}^2_{\nu_u}} := (\mathbb{E}[\int_0^T \|\ell_t\|^2_{\nu_u} dt])^{1/2} < \infty.$$

- $\mathbb{S}^2$ is the set of real-valued RCLL adapted processes ϕ with

$$\|\phi\|_{\mathbb{S}^2} := (\mathbb{E}[\sup_{t\in[0,T]} |\phi_t|^2])^{1/2} < \infty.$$

- $\mathcal{M}\mathbb{H}^2$ is the set of all measurable functions X from I to $\mathbb{H}^2$: $u \mapsto X_u$, satisfying

$$\sup_{u\in I} \|X_u\|^2_{\mathbb{H}^2} = \sup_{u\in I} \mathbb{E}[\int_0^T |X_u(t)|^2 dt] < \infty.$$

  We define $\mathcal{M}L^2(\mathcal{F}_t)$ and $\mathcal{M}\mathbb{S}^2$ similarly.

- $\mathcal{M}\mathbb{H}^2_\nu := (\mathbb{H}^2_{\nu_u})^{\otimes I}$ is the set of all families $\ell := \{\ell_u\}_{u\in I}$ such that

$$\sup_{u\in I}(\mathbb{E}[\int_0^T \|\ell_{u,t}\|^2_{\nu_u} dt])^{1/2} < \infty.$$

- $L^{2,I}(\mathcal{F}_t)$ (for $t \in [0,T]$) is the space of all $\mathcal{F}_t$-measurable family of random variables $X := \{X_u\}_{u\in I}$ satisfying

$$\|X\|_{L^{2,I}} := (\mathbb{E}\big[\int_I |X_u|^2 du\big])^{1/2} < \infty.$$

  We define further the scalar product

$$\langle X, Y \rangle_{L^{2,I}} := \mathbb{E}[\int_I X_u Y_u du].$$

Definition 5.2-Graphon mean field BSDE *A solution of the graphon mean-field BSDE system with jumps* (1.4) *consists of a family of processes* $\Phi := (X_u, Z_u, \ell_u)_{u \in I}$ *with* $(X_u, Z_u, \ell_u) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{H}^2_{\nu_u}$ *for all* u *in* I *satisfying* (1.4)*, where* X_u *is a right continuous with left limit (RCLL)* $\mathbb{R}$*-valued optional process, and* Z_u *(resp.* ℓ_u*) is a* $\mathbb{R}$*-valued predictable process defined on* $\Omega \times [0,T]$ *(resp.* $\Omega \times [0,T] \times E$*) such that the stochastic integral is well defined.*

Assumption 5.1-Intensity measure *For each* $\omega \in [1,2]$*, the function* $I \ni u \mapsto \Phi_u^{-1}(\omega - 1) \in \mathbb{R}$ *is measurable, where* Φ_u *denotes the cumulative distribution function of* ν_u*; we define* $\Phi_u^{-1}(1)$ *as the essential supremum and* $\Phi_u^{-1}(0)$ *as the essential infimum.*

Assumption 5.2-Lipschitz driver *For each* $u \in I$*,*

$$\begin{aligned} &f : \Omega \times [0,T] \times \mathbb{R}^3 \times L^2_{\nu_u} \to \mathbb{R} \\ &(\omega, t, x', x, z, \ell(\cdot)) \mapsto f(\omega, t, x', x, z, \ell(\cdot)) \end{aligned}$$

is $P \otimes \mathcal{B}(\mathbb{R}^3) \otimes \mathcal{B}(L^2_{\nu_u})$ *measurable, and satisfies* $f(\cdot,\cdot,0,0,0,0) \in \mathbb{H}^2$*, and* f *is Lipschitz-continuous in* (x', x, z, ℓ)*, i.e., there exists a constant* $C \geqslant 0$ *such that* $dt \otimes d\mathbb{P}$*-a.s., for each* (x'_1, x_1, z_1, ℓ_1) *and* (x'_2, x_2, z_2, ℓ_2)*, we have*

$$\begin{aligned} &\left| f(\omega, t, x'_1, x_1, z_1, \ell_1(\cdot)) - f(\omega, t, x'_2, x_2, z_2, \ell_2(\cdot)) \right| \\ \leqslant &C(|x'_1 - x'_2| + |x_1 - x_2| + |z_1 - z_2| + \|\ell_1 - \ell_2\|_{\nu_u}). \end{aligned}$$

Define the space

$$\mathcal{M} := \{\Phi_u \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{H}^2_{\nu_u}, \quad \text{for all } u \in I \text{ and satisfying } \|\Phi\|_{\mathcal{M}} < \infty\}.$$

Theorem 5.4-Existence and uniqueness *Let Assumption* 5.1 *and* 5.2 *are satisfied and* $\xi \in \mathcal{M}L^2(\mathcal{F}_T)$*. Then the graphon mean-field BSDE system with jumps* (1.4) *admits a unique solution* $\Phi := (X, Z, \ell) \in \mathcal{M}$*, and* $I \ni u \mapsto \mathcal{L}(X_u)$ *is measurable.*

Assumption 5.3 *We assume that for each* $u \in I$ *and each* $(x', x, z, \ell_1, \ell_2) \in \mathbb{R}^3 \times (L^2_{\nu_u})^2$*, there exists a function* $\phi_{u,t}^{x',x,z,\ell_1,\ell_2} \in L^2_{\nu_u}$ *such that*

$$f(t, x', x, z, \ell_1) - f(t, x', x, z, \ell_2) \geqslant \langle \phi_{u,t}^{x',x,z,\ell_1,\ell_2}, \ell_1 - \ell_2 \rangle_{\nu_u},$$

with

$$\begin{aligned} \phi_{u,t}^{x',x,z,\ell_1,\ell_2} : \quad & [0,T] \times \Omega \times \mathbb{R}^3 \times (L^2_{\nu_u})^2 \mapsto L^2_{\nu_u}; \\ & (t, \omega, x', x, z, \ell_1, \ell_2) \mapsto \phi_{u,t}^{x',x,z,\ell_1,\ell_2}(\omega, \cdot) \end{aligned}$$

$P \otimes \mathcal{B}(\mathbb{R}^3) \otimes \mathcal{B}((L^2_{\nu_u})^2)$ *measurable, bounded and satisfying* $d\mathbb{P} \otimes dt \otimes d\nu_u$ *a.s.*

$$\phi_{u,t}^{x',x,z,\ell_1,\ell_2}(y) \geqslant -1 \quad \text{and} \quad |\phi_{u,t}^{x',x,z,\ell_1,\ell_2}(y)| \leqslant \psi(y),$$

for some $\psi \in L^2_{\nu_u}$*.*

Theorem 5.6-Comparison theorem for graphon mean-field BSDE *Let $\xi^1, \xi^2 \in \mathcal{M}L^2(\mathcal{F}_T)$ and denote by (X^1, Z^1, ℓ^1) and (X^2, Z^2, ℓ^2) the solution of the graphon mean-field BSDE with jumps (1.4) associated to (ξ^1, f_1) and (ξ^2, f_2) respectively. Let f_1 and f_2 both satisfy Assumption 5.2, and further assume that:*

- *At least one of f_1 and f_2 satisfies Assumption 5.3, and the other one (or at least one if both satisfy Assumption 5.3) is non-decreasing in x';*
- *For each $u \in I \backslash H$ with H a zero Lebesgue measure subset of I, $\xi_u^2 \geqslant \xi_u^1$ a.s. and $f_2(\omega, t, x', x, z, \ell) \geqslant f_1(\omega, t, x', x, z, \ell)$ a.s. for all $(t, x', x, z, \ell) \in \mathbb{R}^4 \times L^2_{\nu_u}$.*

Then for all $t \in [0,T]$ and $u \in I \backslash H$, we have $X_u^2(t) \geqslant X_u^1(t)$ almost surely.

Theorem 5.7-Strict comparison for graphon mean-field BSDE *Suppose the assumptions in Theorem 5.6 hold. Further, assume that f_1 satisfies Assumption 5.3 with strict inequality, i.e.,*

$$\phi_{u,t}^{x',x,z,\ell_1,\ell_2}(y) > -1,$$

and $\xi_u^1 \geqslant \xi_u^2$ a.s. for each $u \in I \backslash H$ with H a zero Lebesgue measure subset of I, and $f_1(\omega, t, x', x, z, \ell) \geqslant f_2(\omega, t, x', x, z, \ell)$ a.s. for all $(t, x', x, z, \ell) \in \mathbb{R}^4 \times L^2_{\nu_u}$. Then if $X^1(t_0) = X^2(t_0)$ (i.e., $X_u^1(t_0) = X_u^2(t_0)$ for all $u \in I \backslash H$) for some $t_0 \in [0,T]$, we have $X^1(\cdot) = X^2(\cdot)$ a.s. on $[t_0, T]$, and $f_2(\omega, t, x', x, z, \ell) = f_1(\omega, t, x', x, z, \ell)$ on $[t_0, T]$ for $u \in I \backslash H$.

We next study the convergence result of N-coupled system (1.5) to graphon system (1.4). We need the following assumptions.

Assumption 5.4 *For each $u \in I$,*

(i) $u \to \mathcal{L}(\xi_u)$ is continuous w.r.t. the $\mathcal{W}_2$ metric.

(ii) there exists a finite collection of intervals $\{I_i : i = 1, \ldots, N\}$ such that $I = \cup_i I_i$, and for each $i \in \{1, \ldots, N\}$, we have $G(u,v)$ is continuous at u for each $v \in I \backslash H_i$ for some zero Lebesgue measure set H_i.

Assumption 5.5 *There exists a finite collection of intervals $\{I_i : i = 1, \ldots, N\}$ such that $I = \cup_i I_i$, and for some constant C, we have for all $u_1, u_2 \in I_i$, $v_1, v_2 \in I_j$, and $i, j \in \{1, \ldots, N\}$,*

$$\mathcal{W}_2(\mathcal{L}(\xi_{u_1}), \mathcal{L}(\xi_{u_2})) \leqslant C|u_1 - u_2|,$$

and,

$$|G(u_1, v_1) - G(u_2, v_2)| \leqslant C(|u_1 - u_2| + |v_1 - v_2|).$$

Assumption 5.6-Interaction regularity *For a given graphon G, we say that $\zeta^N := \{\zeta_{ij}^N\}_{i,j \in [N]}$ satisfies regularity assumption with graphon G if either:*

(i) $\zeta_{ij}^N = G(\frac{i}{N}, \frac{j}{N})$;

(ii) $\zeta_{ij}^N = \mathrm{Bernoulli}\big(G(\frac{i}{N}, \frac{j}{N})\big)$ *independently for all* $1 \leqslant i \leqslant j \leqslant N$ *and independent of* $\{W_u, N_u, \xi_u : u \in I\}$.

For notation simplicity, we let all ν_u be a common measure ν. But note that all following results hold for different ν_u. See more details in Chapter 5. The convergence results are as follows.

Theorem 5.14 *Let Assumptions 5.2 and 5.5 be fullfilled. Suppose that* ζ^N *satisfy the regularity assumption 5.6 with graphon* G*, and the terminal conditions* ξ^N *and* ξ *satisfy*

$$\max_{i=1,\dots,N} \mathbb{E}|\xi_i^N - \xi_{\frac{i}{N}}|^2 = O(N^{-1}).$$

Then the unique solutions Φ^N *of* (1.5) *converge to the unique solution of* (1.4) *with the convergence rate* $1/\sqrt{N}$ *and*

$$\begin{aligned}
&\max_{i=1,\dots,N} \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right] \\
&\quad \leqslant CN^{-1} + C \max_{i=1,\dots,N} \mathbb{E}|\xi_i^N - \xi_{\frac{i}{N}}|^2 = O(N^{-1}),
\end{aligned}$$

for all $N \in \mathbb{N}$ *and some constant* C*. Furthermore, for* $\kappa_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i^N(t)}$ *and* $\kappa_t = \int_I \mathcal{L}(X_u(t))du$*, then*

$$\sup_{t\in[0,T]} \mathbb{E}\big[(\mathcal{W}_2(\kappa_t^N, \kappa_t))^2\big] \leqslant CN^{-1/2}.$$

Theorem 5.15 *Let Assumptions 5.2 and 5.5 be fullfilled. Suppose* ζ^N *satisfies the regularity Assumption 5.6 with graphon* G^N*. Then we have*

$$\begin{aligned}
&\max_{i=1,\dots,N} \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right] \\
&\quad \leqslant C\big(\max_{i=1,\dots,N} \mathbb{E}|\xi_i^N - \xi_{\frac{i}{N}}|^2 + \|G - G^N\|_{\infty\to\infty} + N^{-1}\big).
\end{aligned}$$

We introduce the graphon dynamic risk measure.

Definition 5.19 *Let* $T > 0$ *be a time horizon, for a terminal condition* $\xi \in \mathcal{M}L^2(\mathcal{F}_T)$*, we define*

$$\rho_{u,t}(\xi, T) := -X_u(t, \xi, T),$$

for each $u \in I$*, where* $\{X_u(t,\xi,T)\}_{u\in I}$ *is the solution of the graphon mean-field BSDE system* (1.4)*. Then* $\rho_t(\xi,T) := \{\rho_{u,t}(\xi,T)\}_{u\in I}$ *is called the graphon associated dynamic risk measures.*

Define

$$F_u(\omega, t, \mathcal{L}(X_t), x, z, \ell(\cdot)) := \int_I \int_{\mathbb{R}} G(u,y) f(s, x', x, z, \ell(\cdot)) \mu_{y,s}(dx') dy.$$

For each (ω, t) and each $u \in I$, we denote by $(F_u)^*$ the Fenchel-Legendre transform, defined as

$$(F_u)^*(\omega, t, \mathcal{L}(Y), \beta_u, \alpha_u^1, \alpha_u^2) := \sup_{(X,x,z,\ell)\in L^{2,I}(\mathcal{F}_t)\otimes\mathbb{R}^2\otimes L^2_{\nu_u}} \{F_u(\omega, t, \mathcal{L}(X), x, z, \ell) - \langle X, Y\rangle_{L^{2,I}} - \beta_u x - \alpha_u^1 z - \langle \alpha_u^2, \ell_u\rangle_{\nu_u}\}.$$

For given processes (β, γ), we define

$$H_{t,s}^{\beta,\gamma} := \exp\{\int_t^s (\beta_y + \gamma_y) dy\}.$$

We introduce the set $\mathcal{A}_T^I$, which a set of families of processes $(\{\gamma_t^{u,v}\}_{u,v\in I}, \{\beta_{u,t}\}_{u\in I}, \{\alpha_{u,t}\}_{u\in I})_{t\in[0,T]}$ defined in §5.4.2. The dual representation formula for the graphon dynamic risk measure can be characterized by the known features of our graphon mean field BSDE system and is giving as following.

Theorem 5.27 *Suppose f satisfies Assumption 5.2 and 5.3. Moreover, suppose that f is concave with respect to (x', x, z, ℓ) and non-decreasing in x'. Then we have for each $t \in [0,T]$, the expectation of the convex risk-measure ρ_t has the following representation : for each $\xi \in \mathcal{M}L^2(\mathcal{F}_T)$,*

$$\mathbb{E}[\int_I \rho_{v,t}(\xi, T) dv] = \sup_{(\gamma,\beta,\alpha)\in\mathcal{A}_T^I} \{\int_I \mathbb{E}^{\mathcal{Q}_v^\alpha}[-(\int_I H_{t,T}^{\beta_v,\gamma^{u,v}} du)\xi_v] dv - \int_I \zeta_{v,t}(\gamma, \beta, \alpha, T) dv\}, \tag{1.6}$$

where the function ζ, which is called penalty function, defined for each T and $(\gamma, \beta, \alpha) \in \mathcal{A}_T^I$ by

$$\zeta_{v,t}(\gamma,\beta,\alpha,T) = \int_t^T \mathbb{E}^{\mathcal{Q}_v^\alpha}\big[(\int_I H_{t,s}^{\beta_v,\gamma^{u,v}} du)(F_v)^*\big(s, \big(\frac{\Gamma_s^{\alpha_{v_1}} H_{t,s}^{\beta_{v_1},\gamma^{v,v_1}} \gamma_s^{v,v_1}}{\mathbb{E}[\Gamma_s^{\alpha_v} \int_I H_{t,s}^{\beta_v,\gamma^{v_1,v}} dv_1]}\big)_{v_1}, \beta_{v,s}, \alpha_{v,s}^1, \alpha_{v,s}^2(\cdot)\big)\big] ds,$$

with Q_v^α the absolutely continuous probability measure with respect to $\mathbb{P}$ admitting density Γ^{α_v}, which is defined by (5.28) with initial value $\Gamma^{\alpha_{v,t}} = 1$. Moreover, there exists $(\overline{\gamma}, \overline{\beta}, \overline{\alpha}) \in \mathcal{A}_T^I$ attaining the supremum in (1.6). In particular, for each $v \in I$,

$$\mathbb{E}[\rho_{v,t}(\xi, T)] = \mathbb{E}^{\mathcal{Q}^{\overline{\alpha}_v}}[-(\int_I H_{t,T}^{\beta_v,\gamma^{u,v}} du)\xi_v] - \zeta_{v,t}(\overline{\gamma}, \overline{\beta}, \overline{\alpha}, T).$$

1.4.2 Stochastic games with graphon mean field interactions

With increasing interests in graphon interacting systems, there has been a growing research activity on graphon mean field games. Studying stochastic games with heterogeneous interactions or games on networks poses substantial challenges, especially when dealing with games on networks involving a large number of players, since these games may exhibit significant asymmetry. This distinction is particularly noteworthy in the context of games on sparse networks (e.g., [120, 161]). Analyzing games on large networks, especially those with heterogeneous interactions, often relies on tractable limiting (continuum) models. These models provide a convenient approximation for understanding the dynamics of large finite games and offer valuable insights into the complexities of stochastic games in heterogeneously interacting systems.

Chapter 6 aims at developping a graphon interacting model to solve graphon games with heterogeneous interactions and jumps, while maintaining tractability comparable to traditional mean field games (MFGs). The traditional MFG framework is based on a fixed point problem describing the law of the state process $(X(t))_{t\in[0,T]}$ of a typical player. In the graphon game model, we consider a fixed point problem for a family of laws $(X_u(.))_{u\in I}$, which can be viewed as a joint law of (U, X), where X is the randomised state process and the uniform random variable U in $I := [0, 1]$ is interpreted as the "label" variable (order of vertex on network in limiting sense) of the player in the graphon. Despite the heterogeneous interactions, we also include jumps in the dynamics to model the instantaneous impacts. The jumps are induced by Poisson random measures with different intensity measures for different labels, which is a source of individual heterogeneity.

We put our attention to the Markovian feedback controls. The control depends on the current state and its label. Let $\mathcal{A}_I$ be the set of graphon controls α defined as measurable function $\alpha : [0, T] \times I \times \mathbb{R} \to A; (t, u, x) \mapsto \alpha(t, u, x)$, where A is the action set. The dynamics of the controlled graphon system is as follows,

$$\begin{aligned} dX_u^\alpha(s) = &\int_I \int_{\mathbb{R}} G(u,v) b(s, X_u^\alpha(s), x, \alpha(s,u,X_u^\alpha(s)))\mu_{v,t}^\alpha(dx)dvds \\ &+ \int_I \int_{\mathbb{R}} G(u,v)\sigma(s, X_u^\alpha(s), x, \alpha(s,u,X_u^\alpha(s)))\mu_{v,t}^\alpha(dx)dvdW_u(s) \\ &+ \int_E \ell(s, X_u^\alpha(s), e, \alpha(s,u,X_u^\alpha(s)))\widetilde{N}_u(ds,de), \qquad X_u(0) = \xi_u, \qquad u \in I, \end{aligned} \tag{1.7}$$

where $\mu_v^\alpha := \mathcal{L}(X_v^\alpha) \in \mathcal{P}(\mathcal{D})$ and $\mu_{v,s}^\alpha := \mathcal{L}(X_v^\alpha(s)) \in \mathcal{P}(\mathbb{R})$. We **assume** that $\boldsymbol{\xi} := \{\xi_u\}_{u\in I} \in \mathcal{M}L^2(\mathcal{F}_0)$, that is for each $u \in I$, $\xi_u \in L^2(\mathcal{F}_T)$ and the map $u \mapsto \xi_u$ is measurable. The coefficients $b : [0,T] \times \mathbb{R} \times \mathbb{R} \times A \to \mathbb{R}$, $\sigma : [0,T] \times \mathbb{R} \times \mathbb{R} \times A \to \mathbb{R}$ and $\ell : [0,T] \times \mathbb{R} \times E \times A \to \mathbb{R}$ are Lipschitz continuous with respect to all parameters except t. We also assume that σ^2 is bounded from 0.

Notice that in our model, the control term is present not only in the drift, as in [82, 162], but is present also in the diffusion and jump terms. Furthermore, we also have the graphon interaction in the diffusion term, which is not present in the model in [82, 162]. Combined with jumps and controls, more heterogeneity is introduced into our setup, and the interacting dynamic system becomes more complex compared to [47, 55]. Each player with label $u \in I$ seeks to maximize the following objective function:

$$\mathbb{E}\Big[\int_0^T f(t, X_u^\alpha(t), \Lambda\mu_t^\alpha(u), \alpha(t,u,X_u^\alpha(t)))dt + g(X_u^\alpha(T), \Lambda\mu_T^\alpha(u))\Big],$$

where f is some function representing the running cost and g is the cost function at the ending time.

1.4.3 Graphon equilibria and approximate Nash equilibria

The study of graphon mean field games can help to study the finite games on large networks. It is hard to study directly the Nash equilibria of finite games with heterogeneous interaction. Instead, we study it through the graphon games. Let $\mathcal{A}_n$ be the set of measurable functions $\alpha : [0,T] \times \mathbb{R}^n \to A$. The heterogeneous interacting particle system we consider has the following controlled dynamic under control $\{\alpha_i\}_{i\in[n]} \in \mathcal{A}_n^n$,

$$\begin{aligned} dX_i^{(n)}(s) =& \frac{1}{n}\sum_{j=1}^n \zeta_{ij}^{(n)} b(s, X_i^{(n)}(s), X_j^{(n)}(s), \alpha_i(s, X^{(n)}(s)))ds \\ &+ \frac{1}{n}\sum_{j=1}^n \zeta_{ij}^{(n)} \sigma(s, X_i^{(n)}(s), X_j^{(n)}(s), \alpha_i(s, X^{(n)}(s)))dW_i(s) \\ &+ \int_E \ell(s, X_i^{(n)}(s), e, \alpha_i(s, X^{(n)}(s)))\tilde{N}_i(ds, de), \qquad X_i^{(n)}(0) = \xi_i^{(n)}, \end{aligned} \tag{1.8}$$

where $\{W_i, i \in [n]\}$ are i.i.d. Brownian motions, $\{N_i(dt, de), i \in [n]\}$ are independent Poisson random measures, and $\{\xi_i^{(n)}, i \in [n]\}$ are initial conditions. The idea is that when the population size n is large enough, the controlled graphon mean field system (1.7) can be viewed as the limit system of n coupled controlled system (1.8) in some sense. Thus the Nash equilibrium of the finite game should be close to the equilibrium of the graphon game, and hence could be approximated by that of the graphon game. We call such equilibrium of the graphon game the graphon equilibrium and will give the detailed definition. Recent works [36, 82, 162] have studied such approximate Nash equilibria in some special cases. In Chapter 6, we extend the study to a more general framework involving jumps.

We define the concept of graphon equilibrium in our framework and establish its existence by using the compactification method, a powerful technique commonly employed in studying equilibria in various types of mean field games. We also investigate the uniqueness of the graphon equilibrium under certain monotonicity condition. By considering graphon equilibria as benchmarks, we can approximate Nash equilibria in finite games. We use the equilibrium control for graphon games as a benchmark to infer the corresponding control for finite games. Making use of the propagation of chaos results (similar to those in Chapter 5), as the population size grows, the distributions of state processes in finite games converge to those of graphon games under a specific correspondence pattern between the player order $i \in [n]$ and the graphon label $u \in I$. Intuitively, the equilibrium control for each player in the finite game should closely resemble that taken for the corresponding label in the limit graphon system. Therefore, it is natural to select the control associated with the label $\frac{i}{n}$ for the i-th player in an n-player game. When the graphon equilibrium control exhibits continuity with respect to u, we can relax to controls associated with labels close to $\frac{i}{n}$. This approximation follows the principles of classical mean field games theory. However, due to the heterogeneity of interactions in our model, the analysis becomes more intricate. This approximate method for Nash equilibrium applies not only to the model-dependent case but also to the model-free case. It has also found applications in reinforcement learning, as seen in works such as [101, 134].

Contribution of Chapter 6: Stochastic Graphon Mean Field Games with Jumps and Approximate Nash Equilibria

We use the same probabilistic set-up and notation as in Chapter 5.

Assumption 6.1

- *For each* $(t, x, u, \mu) \in [0, T] \times \mathbb{R} \times [0, 1] \times \mathcal{M}^{+}_{\text{Unif}}([0, 1] \times \mathcal{D})$*, there exists* $e \in E$ *such that the set* $K_e[\mu](t, x, u) :=$ $\big\{\big(b(t, x, \Lambda\mu_t(u), a), \sigma^2(t, x, \Lambda\mu_t(u), a), \ell(t, x, e, a), z\big) : a \in A, z \leqslant f(t, x, \Lambda\mu_t(u), a)\big\}$ *is convex.*
- *The map* $e \mapsto \ell(t, x, e, a)$ *is affine for each* $(t, x, a) \in [0, T] \times \mathbb{R} \times A$*.*

In Chapter 6, we also put the same assumption as Assumption 5.2.

Assumption 6.2 *For each* $\omega \in [1, 2]$*, the function* $I \ni u \mapsto \Phi_u^{-1}(\omega - 1) \in \mathbb{R}$ *is measurable, where* Φ_u *denotes the cumulative distribution function of* ν_u*; we define* $\Phi_u^{-1}(1)$ *as the essential supremum and* $\Phi_u^{-1}(0)$ *as the essential infimum.*

For any fixed distribution $\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathcal{D})$ and graphon control $\alpha \in \mathcal{A}_I$, we define the following graphon objective function:

$$J_G(\mu, \alpha) := \mathbb{E}\Big[\int_I \Big(\int_0^T f(t, X_u^\alpha(t), \Lambda\mu_t(u), \alpha(t, u, X_u^\alpha(t)))dt + g(X_u^\alpha(T), \Lambda\mu_T(u))\Big)du\Big], \tag{1.9}$$

where the functions $f : [0, T] \times \mathbb{R} \times \mathcal{M}(\mathbb{R}) \times A \to \mathbb{R}$ and $g : \mathbb{R} \times \mathcal{M}(\mathbb{R}) \to \mathbb{R}$ are bounded continuous w.r.t. all parameters.

Definition 6.2-Graphon equilibrium *A graphon equilibrium is a distribution* $\mu \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathcal{D})$ *such that there exists* $\alpha^\star \in \mathcal{A}_I$ *satisfying*

$$J_G(\mu, \alpha^\star) = \sup_{\alpha \in \mathcal{A}_I} J_G(\mu, \alpha), \quad \textit{with} \quad \mu = \mathcal{L}(X^{\alpha^\star}).$$

Any $\alpha^\star$ *satisfying the above is called an equilibrium control for distribution* μ*.*

Theorem 6.4 and 6.6-Existence and uniqueness of graphon equilibrium *Under Assumption 6.1 and 6.2, there exists at least one graphon equilibrium. Further suppose the following monotonicity condition holds: for each* $a \in A$*, and any* $\mu_1, \mu_2 \in \mathcal{P}_{\text{Unif}}([0, 1] \times \mathbb{R} \times A)$ *and* $t \in [0, T]$*, we have*

$$\int_{[0,1]\times\mathbb{R}} \Big(g(x, \Lambda\bar{\mu}_1(u)) - g(x, \Lambda\bar{\mu}_2(u))\Big)(\bar{\mu}_1 - \bar{\mu}_2)(du, dx) < 0,$$

and

$$\int_{[0,1]\times\mathbb{R}\times A} \Big(f((t, x, \Lambda\bar{\mu}_1(u), a) - f(t, x, \Lambda\bar{\mu}_2(u), a)\Big)(\mu_1 - \mu_2)(du, dx, da) < 0,$$

where $\bar{\mu}$ *is the marginal distribution of the first two coordinates. Then there exists a unique graphon equilibrium.*

Let us define the following gap of objective function between an graphon equilibrium control $\boldsymbol{\alpha}^\star$ and the "optimal" control,

$$\epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) := \sup_{\beta \in \mathcal{A}_n} J_i(\alpha^\star(u_1^{(n)}), \ldots, \alpha^\star(u_{i-1}^{(n)}), \beta, \alpha^\star(u_{i+1}^{(n)}) \ldots, \alpha^\star(u_n^{(n)})) - J_i(\boldsymbol{\alpha}^\star), \tag{1.10}$$

where $\boldsymbol{u}^{(n)} := (u_1^{(n)}, \ldots, u_n^{(n)})$, $\boldsymbol{\alpha}^\star := (\alpha^\star(u_1^{(n)}), \ldots, \alpha^\star(u_n^{(n)}))$ and $\alpha^\star(u_i^{(n)}) := \alpha^\star(\cdot, u_i^{(n)}, \cdot)$, i.e., player i uses the control rule of the graphon equilibrium control of label $u_i^{(n)}$.

Assumption 6.3 *There exists a finite collection of intervals $\{I_i : i = 1, \ldots, n\}$ such that $I = \bigcup_i I_i$ and, for each $i \in \{1, \ldots, n\}$, we have:*

(i) $u \to \mathcal{L}(\xi_u)$ is continuous a.e. on I_i w.r.t. the $\mathcal{W}_2$ metric.

(ii) For each $j \in \{1, \ldots, n\}$, $G(u, v)$ is continuous in u and v a.e. on $I_i \times I_j$.

(iii) The intensity measure ν_u is continuous in u for the Wasserstein distance $\mathcal{W}_2$ on each I_i.

Assumption 6.4 *There exists a finite collection of intervals $\{I_i : i = 1, \ldots, n\}$ such that $I = \bigcup_i I_i$, and for some constant C, we have for all $u_1, u_2 \in I_i$, $v_1, v_2 \in I_j$, and $i, j \in \{1, \ldots, N\}$,*

$$\mathcal{W}_2(\mathcal{L}(\xi_{u_1}), \mathcal{L}(\xi_{u_2})) \leqslant C|u_1 - u_2|,$$

$$|G(u_1, v_1) - G(u_2, v_2)| \leqslant C(|u_1 - u_2| + |v_1 - v_2|),$$

and

$$\mathcal{W}_2(\nu_{u_1}, \nu_{u_2}) \leqslant C|u_1 - u_2|.$$

Assumption 6.5-Interaction regularity *We say $\zeta^{(n)} := \{\zeta_{ij}^{(n)}\}_{i,j \in [n]}$ satisfies the regularity assumption with graphon G if either:*

(i) $\zeta_{ij}^{(n)} = G(\frac{i}{n}, \frac{j}{n})$;

(ii) $\zeta_{ij}^{(n)} = \mathrm{Bernoulli}\Big(G(\frac{i}{n}, \frac{j}{n})\Big)$ independently for all $1 \leqslant i \leqslant j \leqslant n$ and independent of $\{W_u, N_u, \xi_u : u \in I\}$ and $\{W_i, N_i, \xi_i : i \in [n]\}$.

Let $\|\cdot\|_{\mathbb{S}_T^2} := \sup_{s \in [0,T]} \mathbb{E}|\cdot_s|^2$. The propagation result is as follows.

Theorem 6.10-Large population convergence *Let $\alpha(t, u, x)$ be a Lipschitz function on (u, x), and let $\alpha_i^{(n)}(t, x) = \alpha(t, \frac{i}{n}, x)$. Let $X^{(n)}$ and X be the solutions of (1.8) and (1.7) respectively, with initial conditions $\xi^{(n)}$ and ξ, controls $\alpha^{(n)} := (\alpha_i^{(n)})_{i \in [n]}$ and α. Suppose Assumption 6.4 holds with G, and $\zeta^{(n)}$ satisfies the regularity Assumption 6.5 with G_n, where $\{G_n\}_n$ is a sequence of step graphons such that $\|G - G_n\|_\square \to 0$. Then we have the following convergence result for the empirical mean of the neighborhood measure (defined in (6.9)):*

$$\frac{1}{n} \sum_{i=1}^n M_i^{(n)} \to \int_I \Lambda\mu(v) dv,$$

in probability in the weak sense, where $\mu := \mathcal{L}(X)$. Furthermore, for each $i \in [n]$ and any Lipschitz continuous bounded function h from $\mathcal{D}$, we have (for some constant $C > 0$)

$$\mathbb{E}\Big[\langle h, M_i^{(n)}\rangle - \langle h, \Lambda\mu(\frac{i}{n})\rangle\Big]^2 \leqslant \frac{C}{n} \sum_{j=1}^n \mathbb{E}\Big|\xi_j^{(n)} - \xi_{\frac{j}{n}}\Big|^2 + C\|G_n - G\|_\square + C\|G_n - G\|_{\infty\to\infty} + \frac{C}{n}.$$

If W_i, N_i and $W_{\frac{i}{n}}, N_{\frac{i}{n}}$ are the same for each $i \in [n]$, then we have

$$\frac{1}{n}\sum_{i=1}^{n}\|X_i^{(n)} - X_{\frac{i}{n}}\|_{\mathbb{S}_T^2}^2 \leqslant C\Big(\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_{\Box} + \frac{1}{n}\Big),$$

and moreover

$$\max_{i\in[n]}\|X_i^{(n)} - X_{\frac{i}{n}}\|_{\mathbb{S}_T^2}^2 \leqslant C\Big(\max_{i\in[n]}\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_{\infty\to\infty} + \frac{1}{n}\Big).$$

Assumption 6.6-Concavity

- *$f(t,x,\mu,a)$ is concave in x and strictly concave in a.*
- *For all $\lambda \in [0,1]$, $a_1, a_2 \in A$,*

  $$\lambda f(t,x,\mu,a_1) + (1-\lambda)f(t,x,\mu,a_2) \leqslant f(t,x,\mu,\bar{a}_\lambda),$$

  where $\bar{a}_\lambda = \bar{a}_\lambda(t,x,\mu)$ is the solution to

  $$b(t,x,\mu,\bar{a}_\lambda) = \lambda b(t,x,\mu,a_1) + (1-\lambda)b(t,x,\mu,a_2).$$

Lemma 6.17-Stability of control *Suppose that Assumptions 6.3 and 6.6 are satisfied. Then there exists a unique optimizer $\alpha_u^\star$ for $\sup_{\alpha\in\mathcal{A}_n} J_G^{u,\xi_u}(\mu,\alpha)$. Let $\alpha^\star(t,u,x) := \alpha_u^\star(t,x)$. We have $\alpha^\star(t,u,x)$ is (piecewise) continuous in (u,x), and the law of $X_u^{\alpha^\star}$ is (piecewise) continuous in u in the weak sense. Furthermore, if G, f, g are all Lipschitz continuous and Assumption 6.4 is satisfied, then all the continuities become Lipschitz continuities.*

For each $i \in [n]$, we define $\mathcal{I}_i^{(n)} := (\partial_- I_j, \frac{i}{n}]$ if $\frac{i-1}{n} \notin I_j$, $\frac{i}{n} \in I_j$; $\mathcal{I}_i^{(n)} := \big(\frac{i-1}{n}, \frac{i}{n}\big]$ if $\frac{i-1}{n}, \frac{i}{n} \in I_j$; $\mathcal{I}_i^{(n)} := \big[\frac{i-1}{n}, \partial_+ I_j\big)$ if $\frac{i}{n} \in I_j$ and $\frac{i+1}{n} \notin I_j$, where ∂_- and ∂_+ denote the lower and upper borders, respectively. We call $\{G_n\}_{n\in\mathbb{N}}$ a sequence of step graphons if, for each $n \in \mathbb{N}$, G_n is a graphon and satisfies $G_n(u,v) = G_n\Big(\frac{[nu]}{n}, \frac{[nv]}{n}\Big)$ for all $(u,v) \in I \times I$.

The approximate Nash equilibria results are the followings.

Theorem 6.15, 6.18 and 6.20 *We have the following approximate Nash equilibria for four different types of graphon under different conditions :*

- **Piecewise constant graphon.** *Suppose $\zeta^{(n)}$ satisfies regularity Assumption 6.5 with G and Assumption 6.3 (i) holds. If*

  $$\max_{i=1,\dots,n}\mathbb{E}\big|\xi_i^{(n)} - \xi_{\frac{i}{n}}\big|^2 = O(n^{-1}),$$

  then taking $u_i^{(n)} = \frac{i}{n}$, we have as $n \to \infty$,

  $$\max_{i=1,\dots,n}\epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \to 0.$$

Moreover, if the initial condition is Lipschitz, satisfying (6.10)*, then we have*

$$\max_{i=1,\ldots,n} \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) = O(n^{-1}).$$

- **Continuous graphon.** *Suppose Assumption 6.6 holds, $\zeta^{(n)}$ satisfies the regularity Assumption 6.5 with step graphon G_n, and $\|G - G_n\|_\square \to 0$. Suppose Assumption 6.3 holds, G is continuous, and the initial condition satisfies $\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 \to 0$. Then we have*

$$\text{ess sup}_{\boldsymbol{u}^{(n)} \in \mathcal{I}_1^{(n)} \times \cdots \times \mathcal{I}_n^{(n)}} \frac{1}{n} \sum_{i=1}^n \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \to 0.$$

  Furthermore, if $\|G - G_n\|_{\infty\to\infty} \to 0$ and $\max_{i=1,\ldots,n} \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 \to 0$, then we have

$$\text{ess sup}_{\boldsymbol{u}^{(n)} \in \mathcal{I}_1^{(n)} \times \cdots \times \mathcal{I}_n^{(n)}} \max_{i=1,\ldots,n} \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \to 0.$$

- **Lipschitz Continuous graphon.** *Suppose Assumption 6.6 holds, $\zeta^{(n)}$ satisfies the regularity Assumption 6.5 with step graphon G_n, and $\|G - G_n\|_\square \to 0$. Suppose Assumption 6.4 holds, G, f, and g are Lipschitz continuous, and the initial condition satisfies $\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 = O(n^{-1})$. Then we have*

$$\text{ess sup}_{\boldsymbol{u}^{(n)} \in \mathcal{I}_1^{(n)} \times \cdots \times \mathcal{I}_n^{(n)}} \frac{1}{n} \sum_{i=1}^n \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) = O(n^{-1}).$$

  Furthermore, if $\|G - G_n\|_{\infty\to\infty} \to 0$ and $\max_{i=1,\ldots,n} \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 = O(n^{-1})$, then we have

$$\text{ess sup}_{\boldsymbol{u}^{(n)} \in \mathcal{I}_1^{(n)} \times \cdots \times \mathcal{I}_n^{(n)}} \max_{i=1,\ldots,n} \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) = O(n^{-1}).$$

- **Sampling graphon.** *Suppose Assumption 6.6 and 6.3 hold. Let $\zeta^{(n)}$ be sampled from the continuous graphon G. If the initial condition satisfies $\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 \to 0$, then we have, for both ways of sampling defined in Section 6.5, as $n \to \infty$,*

$$\text{ess sup}_{\boldsymbol{u}^{(n)} \in \mathcal{I}_1^{(n)} \times \cdots \times \mathcal{I}_n^{(n)}} \frac{1}{n} \sum_{i=1}^n \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \to 0.$$

1.5 Publications and working papers

1. Hamed Amini, Zhongyuan Cao, and Agnès Sulem. Limit theorems for default contagion and systemic risk. *Available at SSRN 3811107 and in revision for Mathematics of Operations Research*, 2021. https://ssrn.com/abstract=3811107

2. Hamed Amini, Zhongyuan Cao, and Agnès Sulem. Fire sales, default cascades and complex financial networks. *Available at SSRN 3935450 and in revision for Mathematical Finance*, 2021. https://ssrn.com/abstract=3935450

3. Hamed Amini, Zhongyuan Cao, and Agnès Sulem. The default cascade process in stochastic financial networks. *Available at SSRN 4020598*, 2021. https://ssrn.com/abstract=4020598

4. Hamed Amini, Zhongyuan Cao, and Agnès Sulem. Graphon mean-field backward stochastic differential equations with jumps and associated dynamic risk measures. *Available at SSRN 4162616*, 2022. https://ssrn.com/abstract=4162616

5. Hamed Amini, Zhongyuan Cao, Andreea Minca, and Agnès Sulem. Ruin probabilities for risk processes in stochastic networks. *Available at SSRN 4355988*, 2022. https://ssrn.com/abstract=4355988

6. Hamed Amini, Zhongyuan Cao, and Agnès Sulem. Stochastic graphon mean field games with jumps and approximate nash equilibria. *Available at SSRN 4412999*, 2023. https://ssrn.com/abstract=4412999

Part I

SYSTEMIC RISK MODELS IN COMPLEX FINANCIAL NETWORKS

Chapter 2

Limit Theorems for Default Contagion and Systemic Risk

This chapter is based on papers [1] and [3] in the publication list of Section 1.5.

Abstract. We consider a general tractable model for default contagion and systemic risk in a heterogeneous financial network, subject to an exogenous macroeconomic shock. We show that, under some regularity assumptions, the default cascade model can be transferred to a death process problem represented by a balls-and-bins model. We state various limit theorems regarding the final size of the default cascade. Under suitable assumptions on the degree and threshold distributions, we prove that the final size of default cascade has asymptotically Gaussian fluctuations. We next state limit theorems for different system-wide wealth aggregation functions, which allow us to provide systemic risk measures in relation with the structure and heterogeneity of the financial network. We finally show how these results can be used by a social planner to optimally target interventions during a financial crisis, with a budget constraint and under partial information of the financial network. Furthermore, we also study the default cascade processes in stochastic networks and obtain limit theorems.

Keywords: Systemic Risk, Default Contagion, Financial Networks, Random Graphs.

2.1 Introduction

The financial crisis 2007-2009 has illustrated the significance of network structure on the amplification of initial shocks in the banking system to the level of the global financial system, leading to an economic recession. An important literature on systemic risk and financial networks has emerged, see e.g. [91, 146] for two recent surveys and references there.

Chapter 2 studies structural and dynamic models for loss propagation in the network of liabilities. This is in contrast to the well-known systemic risk indicators such as CoVAR [192] or SES [2], which are based on measuring losses in terms of market equity. Empirical studies on network topology of banking systems reveal very different structures; from centralized networks as in [176] to core-periphery structures [99, 122, 166] and scale-free structures as in [69, 93]. The main objective of this chapter is to provide limit theorems and use them to establish a link between the (final) size of default cascade and the structure and heterogeneity of financial networks. This chapter also studies limit theorems to quantify the system-wide wealth and systemic risk in the financial system.

A crucial point in systemic risk modeling is the available information. As pointed out in [30, 126, 144, 187, 193], only partial information is, in general, available on the financial network, e.g., the total size of the assets and liabilities for each institution. Our probabilistic approach allows us to deal with an incomplete observation of the system connections. We reduce the dimension of the problem by considering a classification of financial institutions according to different types (characteristics), in an appropriate type space $\mathcal{X}$. Our limit theorems relate the fraction of defaults to "averaged" quantities concerning types/degrees and their propensity to default (the fraction of each type and the threshold distribution for each type) rather than requiring knowledge of the strength of individual exposures. The heterogeneity of exposures are encoded in the type dependent threshold distributions.

An extensive research in systemic risk and financial networks focuses on equilibrium approach, to derive recovery rates from a fixed point equation [111, 114, 131, 185]. This relies on the assumption that all debts are instantaneously cleared, unlikely to hold in reality. Even in a given shock scenario, recovery rates are uncertain. For example, recovery rates after the failure of Lehman were around 8% ([175]). In this chapter, we model recovery rates as given. The model could be easily extended to a setup with random recovery rates satisfying some cash-flow consistency conditions, see e.g. [24].

Our work is related to the literature on network structure and threshold models of contagion, see e.g., [157, 165, 174, 200] in the context of (undirected) social networks. As shown in [7, 28, 105, 124, 179], network topology plays an important role for default propagation in financial systems. In particular, [1] compares regular financial networks, and shows that the completely connected system is the most stable for small shocks but the least stable for large shocks (and vice-versa for the ring network). In [21], the authors present a more general framework to find the optimal network structure for reducing the systemic risk. Recent papers, see e.g., [43, 116], consider the endogenous formation of financial networks.

The primary innovations and results of this chapter are in multiple directions.

First, we generalize the default contagion model of [20] and allow for more network heterogeneity by considering the type-dependent threshold model. These types may be calibrated to real-world data by using machine learning techniques for classification. We transfer the default cascade model to a death process problem represented by a balls-and-bins model [1]. This allows us to provide limit theorems for the dynamic contagion model [2].

Second and more importantly, it is the first (to the best of our knowledge), to provide central limit theorems for default contagion and systemic risk in random financial networks. Related to our work, [20] derives a law of large number for default contagion in configuration model and provides a criterion for the resilience of a financial network to insolvency contagion, based on the connectivity and the structure of contagious links (i.e., those exposures of a bank larger than its capital). Here, we show that the final size of default cascade has asymptotically Gaussian fluctuations and state various theorems regarding the joint asymptotic normality between different contagion parameters, including the number of solvent banks, defaulted banks, healthy links (those initiated by solvent banks) and infected links (those initiated by defaulted banks) at any time t. We use Monte Carlo methods to investigate systems with finite number of institutions and compare them with our central limit theorems. We show how our limit theorems can be used to construct confidence intervals for the size of contagion.

Third, we provide limit theorems for system-wide wealth aggregation functions, which can be used for measuring and quantifying systemic risk. This also provides an indicator for the health of financial system in different stress scenarios.

Finally, we consider a social planner who seeks to optimally target interventions during a financial crisis, under partial information of the financial network and with a budget constraint. We show how limit theorems allow us to simplify the optimization problem. The complete information setup has been recently studied in [125, 145].

Aside from the application to default contagion and systemic risk in financial networks, our results contribute to the literature on diffusion processes on random graphs. Related problems are the k-core and bootstrap percolation. The k-core of any finite graph can be found by removing nodes with degree less than k, in any order, until no such nodes exist. The asymptotic normality of k-core has been studied in [152]. The bootstrap percolation is a diffusion process that has been studied on a variety of graphs, see e.g., [9, 10, 148]. In bootstrap percolation process, for a fixed threshold $\theta \geqslant 2$, there is an initially subset of active nodes and in each round, each inactive node that has at least θ active neighbors becomes active and remains so forever. The asymptotic normality of bootstrap percolation has been recently studied in [13]. Our results generalize those of previous studies on bootstrap percolation and k-core in random graphs to the case of heterogeneous random directed networks with type-dependent random thresholds.

Our proof of central limit theorem is a direct generalization of [13, 152] with significantly more involved calculations. The key idea in the proof of [152] is to transfer the (k-core) process to a death process problem represented by a balls-and-bins model. After that we appeal to a martingale limit

[1] The balls-and-bins model has been previously used in the economic literature; see e.g. [34] for a balls-and-bins model of international trade.

[2] Although this chapter does not study the dynamic case, this virtual time (associated to the corresponding death process) allows us to study the equilibrium and the final state of contagion. In Chapter 2, we show how the time-change technique for Markov processes (see e.g., [184, III. (21.7)]) can be used to apply these limit theorems to other Markovian *dynamic* default cascade processes.

theorem from [147] to derive the limiting distributions. However, our model is more general and new difficulties arise in treating the Markov process and proving the convergence results. For each institution, since the interbank losses are random, the default threshold is also random. This makes the covariances much harder to calculate and some convergence conditions become harder to verify. Besides, our financial system is also directed and we need to divide the half-edges into two types (out and in).

We end this introduction by the following remark. In the real world application, using limit theorems requires some caution. For example, in order to let the asymptotic analysis to be relevant, the financial network should be sufficiently large (see Figure 2.2). This could be true for example at the level of a large economic zone. Moreover, financial networks may have small cycles. Most existing literature on random networks features locally tree-like property. However, recent literature shows that the basic configuration model can be extended to incorporate clustering; see e.g., [98, 195]. Moreover, following the recent literature on portfolio compression in financial networks (see e.g., [21, 104, 197]), the study of default contagion and systemic risk in sparse financial networks regime becomes significantly important, as portfolio compression removes small cycles. In light of its tractability and interpretability, as well as its potential to be enriched with clustering, in this chapter we use the configuration model as our base model. Note that the closed form interpretable limit theorems that we provide could also serve as a mandate for regulators to collect data on those specific network characteristics and assess systemic risk via more intensive computational methods.

Outline. The chapter is organized as follows. Section 2.2 introduces a model for the network of financial counterparties and describes a mechanism for default cascade in such a network, after an exogenous macroeconomic shock. We also describe how the default contagion model can be transferred to a death process problem represented by a balls-and-bins model. Section 2.3 gives our main results on limit theorems for the final size of default cascade. In particular, under some regularity assumptions, we show that different default contagion parameters have asymptotically Gaussian fluctuations. Section 2.4 states limit theorems for different financial system aggregation functions, which are used for measuring and quantifying systemic risk. Section 2.5 shows how these limit theorems can be used by a social planner to optimally target interventions during a financial crisis, with a budget constraint and under partial information of the financial network. Proof of main theorems are given in Section 2.7. Section 2.9 concludes. Proof of lemmas are provided in Appendix.

Notation. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If $c \in \mathbb{R}$, we write $X_n \xrightarrow{p} c$ to denote that X_n converges in probability to c, that is, for any $\epsilon > 0$, we have $\mathbb{P}(|X_n - c| > \epsilon) \to 0$ as $n \to \infty$. We write $X_n \xrightarrow{d} X$ to denote that X_n converges in distribution to X. Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers that tends to infinity as $n \to \infty$. We write $X_n = o_p(a_n)$, if $|X_n|/a_n \xrightarrow{p} 0$. If $\mathcal{E}_n$ is a measurable subset of Ω, for any $n \in \mathbb{N}$, we say that the sequence $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$ occurs with high probability (w.h.p.) or almost surely (a.s.) if $\mathbb{P}(\mathcal{E}_n) = 1 - o(1)$, as $n \to \infty$. Also, we denote by $\mathsf{Bin}(k, p)$ a binomial distribution corresponding to the number of successes of a sequence of k independent Bernoulli trials each having probability of success p. We denote by $\mathcal{D}[0, \infty)$ the standard space of right-continuous functions with left limits on $[0, \infty)$

equipped with the Skorokhod topology (see e.g. [147, 156]).

We suppress the dependence of parameters on the size of the network n when it is clear in the context.

2.2 Model

2.2.1 Financial Network and Default Cascade

Consider an economy $\mathcal{E}_n$ consisting of n interlinked financial institutions (banks) denoted by $[n] := \{1, 2, \ldots, n\}$ that intermediate credit among end-users. Banks hold claims on each other. Interbank liabilities are represented by a matrix of nominal liabilities (ℓ_{ij}). For two financial institutions $i, j \in [n]$, $\ell_{ij} \geqslant 0$ denotes the cash-amount that bank i owes bank j. This also represents the maximum loss related to direct claims, incurred by bank j upon the default of bank i. The total nominal liabilities of bank i sum up to $\ell_i = \sum_{j\in[n]} \ell_{ij}$, while the total value of interbank assets sum up to $a_i = \sum_{j\in[n]} \ell_{ji}$. The total value of claims held by end-users on bank i (deposits) is given by d_i. The total value of claims held by bank i on end-users (external assets) is denoted by e_i. In a stress testing framework, we apply a (fractional) shock ϵ_i to the external assets of bank i. The capital of bank i after the shock denoted by $c_i = c_i(\epsilon_i)$ satisfies $c_i = (1 - \epsilon_i)e_i + a_i - \ell_i - d_i$, which represents the capacity of bank i to absorb losses while remaining solvent. A financial institution $i \in [n]$ is said to be *fundamentally insolvent* if its capital after the shock is negative, i.e. $c_i < 0$. For a given shock scenario $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n) \in [0, 1]^n$, we define the set of fundamental defaults $\mathcal{D}_0(\boldsymbol{\epsilon}) = \{i \in [n] : c_i(\epsilon_i) < 0\}$. Following the fundamentally insolvent institutions $\mathcal{D}_0(\boldsymbol{\epsilon})$, there will be a default contagion process. Let us denote by $R_{ij} = R_{ij}(\boldsymbol{\epsilon})$ the recovery rate of the liability of i to j, and by $\mathcal{R} = (R_{ij})$ the matrix of recovery rates. Since any bank i cannot pay more than its external assets $(1 - \epsilon_i)e_i$ plus what it recovered from its debtors, the recovery rates of i should satisfy the following cash-flow consistency constraints

$$(1 - \epsilon_i)e_i + \sum_{j=1}^{n} R_{ji}\ell_{ji} \geqslant \sum_{j=1}^{n} R_{ij}\ell_{ij} + d_i.$$

Given the shock scenario $\boldsymbol{\epsilon}$ and the matrix of recovery rates $\mathcal{R}$, following the set of fundamental default $\mathcal{D}_0$, there is a default cascade that reaches the set $\mathcal{D}^\star$ in equilibrium. This represents the set of financial institutions whose capital is insufficient to absorb losses and should satisfy the following fixed point equation: $\mathcal{D}^\star = \mathcal{D}^\star(\boldsymbol{\epsilon}, \mathcal{R}) = \Big\{i \in [n] : c_i(\epsilon_i) < \sum_{j\in\mathcal{D}^\star}(1 - R_{ji})\ell_{ji}\Big\}$. As stated in [24], the above fixed point default cascade set has in general multiple solutions. The smallest fixed point set which corresponds to smallest number of defaults can be obtained by starting from $\mathcal{D}_0$ and setting at step k: $\mathcal{D}_k = \mathcal{D}_k(\boldsymbol{\epsilon}, \mathcal{R}) = \big\{i \in [n] : c_i(\epsilon_i) < \sum_{j\in\mathcal{D}_{k-1}}(1 - R_{ji})\ell_{ji}\big\}$.

The cascade ends at the first time k such that $\mathcal{D}_k = \mathcal{D}_{k-1}$. Hence in a financial network of size n, the cascade will end after at most $n - 1$ steps and $\mathcal{D}_{n-1} = \mathcal{D}_{n-1}(\boldsymbol{\epsilon}, \mathcal{R})$ represents the final set of insolvent institutions starting from the initial set of defaults $\mathcal{D}_0$.

2.2.2 Node Classification and Configuration Model

In the following, in order to reduce the dimensionality of the problem, we consider a classification of financial institutions into a countable (finite or infinite) possible set of characteristics $\mathcal{X}$. All (observable) characteristics for institution i are encoded in $x_i = (d_i^+, d_i^-, t_i, ...) \in \mathcal{X}$, where d_i^+ denotes the in-degree (number of institutions i is exposed to), d_i^- denotes the out-degree (number of institutions exposed to i) and t_i denotes any other institution's type specific (e.g., credit rating, seniority class, etc.). As we are interested in limit theorems, we consider a sequence of economies $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$, indexed by the number of institutions. In particular, in the economy $\mathcal{E}_n$, the characteristic of any institution $i \in [n]$ is denoted by $x_i^{(n)} = (d_i^{+(n)}, d_i^{-(n)}, t_i^{(n)}, ...) \in \mathcal{X}$. Without loss of generality, the institutions in the same class $x \in \mathcal{X}$ are assumed to have the same number of creditors (denoted by d_x^-), the same number of debtors (denoted by d_x^+) and the same other features.

Under some regularity assumptions detailed below, one can show that the information regarding assets, liabilities, capital after shocks and recovery rates could all be encoded in a single threshold distribution function (see [20] for a similar setup). Namely, for a given shock scenario $\boldsymbol{\epsilon}$ and matrix of recovery rates $\mathcal{R}$[3], we introduce, for any institution $i \in [n]$, the (random) threshold $\Theta_i = \Theta_i^{(n)}$ which measures the number of defaults i can tolerate before becoming insolvent, if its counterparties default in a uniformly at random order, i.e., when i's debtors default order environment is chosen uniformly at random among all possible permutations. Let us denote by $\Sigma_i^{(n)}$ the set of all permutations of counterparties of institution i, i.e. the set $\{j \in [n] | \ell_{ji} > 0\}$. Mathematically, in a permutation environment $\sigma_i \in \Sigma_i^{(n)}$, the default threshold of institution i which belongs to type x is defined as

$$\Theta_i^{(n)}(\sigma_i; (\ell_{ij})) := \min\{k \geqslant 0 | c_i(\boldsymbol{\epsilon}) \leqslant \sum_{j=1}^{k} (1 - R_{\sigma_i(j)i}) \ell_{\sigma_i(j)i}\}.$$

Then with fixed liability matrix (ℓ_{ij}), the probability that an type x institution admits threshold θ is given by

$$q_x^{(n)}(\theta; (\ell_{ij})) := \frac{\#\{(i, \sigma_i) | x_i^{(n)} = x, \sigma_i \in \Sigma_i^{(n)}, \Theta_i^{(n)}(\sigma_i; (\ell_{ij})) = \theta\}}{n\mu_x^{(n)} d_x^+!}.$$

For tractability, we make the following assumption on the probability threshold functions.

Assumption 2.1. *We assume that there exists a classification of the financial institutions into a countable set of possible characteristics $\mathcal{X}$ such that, for each $n \in \mathbb{N}$, the institutions in the same characteristic class have the same threshold distribution function (denoted by $q_x^{(n)}$ for institutions in class $x \in \mathcal{X}$). Namely, for economy $\mathcal{E}_n, i \in [n]$ and for all $\theta \in \mathbb{N}$: $\mathbb{P}(\Theta_i^{(n)} = \theta) = q_{x_i^{(n)}}^{(n)}(\theta)$.*

In particular, in the network of size n, $q_x^{(n)}(0)$ represents the proportion of initially insolvent institutions with type $x \in \mathcal{X}$. As discussed in [16, 20], this assumption is fulfilled e.g. for independent (type-dependent) random losses. For $x \in \mathcal{X}$, let ϵ_x be a random variable and $\{L_x^{(k)}\}_{k=1}^{\infty}$ be a set of

[3] Our results can be extended to a framework with independent random recovery rates; see e.g. [24] for a discussion.

i.i.d. positive continuous random variables. Let c_x^{in} be the inter-network capital buffer aside from the outside assets. Then the limit threshold distribution can be characterized as

$$q_x(0) := \mathbb{P}\big(e_x(1-\epsilon_x) + c_x^{in} \leqslant 0\big),$$

$$q_x(1) := \mathbb{P}\big(0 < e_x(1-\epsilon_x) + c_x^{in} \leqslant L_x^{(1)}\big),$$

and for all $\theta \geqslant 2$,

$$q_x(\theta) := \mathbb{P}\big(L_x^{(1)} + \cdots + L_x^{(\theta-1)} < e_x(1-\epsilon_x) + c_x^{in} \leqslant L_x^{(1)} + \cdots + L_x^{(\theta)}\big).$$

Let $\mu_x^{(n)}$ denote the fraction of institutions with characteristic $x \in \mathcal{X}$ in the economy $\mathcal{E}_n$. In order to study the asymptotics, it is natural to assume the following.

Assumption 2.2. *We assume that for some probability distribution functions μ and q over the set of characteristics $\mathcal{X}$ and independent of n, we have $\mu_x^{(n)} \to \mu_x$ and $q_x^{(n)}(\theta) \to q_x(\theta)$ as $n \to \infty$, for all $x \in \mathcal{X}$ and $\theta = 0, 1, \ldots, d_x^+$. Moreover, we assume that $\sum_{\theta=0}^{d_x^+} q_x(\theta) = 1$ for all $x \in \mathcal{X}$.*

Note that, for simplicity of the computations, the threshold distributions are assumed to satisfy $\sum_{\theta=0}^{d_x^+} q_x(\theta) = 1$ for all $x \in \mathcal{X}$. One can define $q_x(\infty) := 1 - \sum_{\theta=0}^{d_x^+} q_x(\theta)$ and generalize all results with slight changes.

Given the degree sequences $\mathbf{d}_n^+ = (d_1^+, \ldots, d_n^+)$ and $\mathbf{d}_n^- = (d_1^-, \ldots, d_n^-)$ such that $\sum_{i\in[n]} d_i^+ = \sum_{i\in[n]} d_i^-$, we associate to each institution i two sets: $\mathcal{H}_i^+$ the set of incoming half-edges and $\mathcal{H}_i^-$ the set of outgoing half-edges, with $|\mathcal{H}_i^+| = d_i^+$ and $|\mathcal{H}_i^-| = d_i^-$. Let $\mathbb{H}^+ = \bigcup_{i=1}^n \mathcal{H}_i^+$ and $\mathbb{H}^- = \bigcup_{i=1}^n \mathcal{H}_i^-$. A *configuration* is a matching of $\mathbb{H}^+$ with $\mathbb{H}^-$. When an out-going half-edge of instituion i is matched with an in-coming half-edge of institution j, a directed edge from i to j appears in the graph. The configuration model is the random directed multigraph which is uniformly distributed across all configurations. The random graph constructed by the configuration model is denoted by $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$. It is then easy to show that conditioned on the multigraph being a simple graph, we obtain a uniformly distributed random graph with these given degree sequences denoted by $\mathcal{G}_*^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$. In particular, any property which holds with high probability on the configuration model also holds with high probability conditioned on this random graph being simple (for the random graph $\mathcal{G}_*^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$) provided $\liminf_{n\to\infty} \mathbb{P}(\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-) \text{ simple}) > 0$, see e.g. [194].

2.2.3 Death Process and Final Solvent Institutions

We consider the default contagion process in the random financial network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$, initiated by the set of fundamentally insolvent institutions $\mathcal{D}_0$. Recall that Θ_i denotes the random threshold of institution $i \in [n]$ which measures the number of defaults i can tolerate before becoming insolvent in the uniformly chosen i's counterparties default order environment. By Assumption 2.1 and standard coupling arguments, as also proved in [20], we assume that these thresholds are assigned initially to any institution $i \in [n]$ according to the distribution $q_{x_i}^{(n)}(\cdot)$.

Finding the final solvent institutions. We consider the above default contagion progress in the following way. At time 0 in the (random) graph $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$, all institutions with threshold 0 become defaulted. We remove all the initially defaulted institutions $\mathcal{D}_0$ from the network. Next, in order to find $\mathcal{D}_1$, we identify the partners of $\mathcal{D}_0$. Note that the out-degree and in-degree of each institution in the network induced by $[n] \backslash \mathcal{D}_0$ is less than or equal to those in the previous network. At step $k \in \mathbb{N}$, the default set $\mathcal{D}_k$ can be identified by

$$\mathcal{D}_k = \Big\{ i \in [n] : \sum_{j:j \to i} \mathbb{1}\{j \in \mathcal{D}_{k-1}\} \geqslant \Theta_i \Big\}, \tag{2.1}$$

where $\mathbb{1}\{\mathcal{E}\}$ denotes the indicator of an event $\mathcal{E}$, i.e., this is 1 if $\mathcal{E}$ holds and 0 otherwise. We denote the in-degree and out-degree of each institution i after k steps evolution by $d_i^+(k)$ and $d_i^-(k)$ respectively. Note that initially $d_i^+(0) = d_i^+$ and $d_i^-(0) = d_i^-$. At step k, we remove all institutions $i \in [n]$ with $d_i^+(k) < d_i^+ - \Theta_i$. At the end of the above procedure, all the removed institutions are defaulted and the remaining institutions are solvent.

Transferring to a death process problem represented by balls-and-bins. It is not hard to see that the calendar time does not take any important role in the contagion process. We can define the time interval as we want. So instead of removing institutions, we can also remove the links and define a proper time interval between two successive removals. Namely, at each step, we only look at one removal (interaction) between two institutions, yielding at most one default. In the following, we simultaneously run the default contagion process and construct the configuration model. We call all out half-edges and in half-edges that belong to a defaulted (solvent) institution the *infected* (*healthy*) half-edges. We consider all the institutions as bins and all the (in and out) half-edges as (in and out) balls. Consequently, the bins are called defaulted (**D** type) or solvent (**S** type) according to their states as institutions. Similarly the balls are called infected (**I** type) or healthy ($\mathbb{H}$ type) when they are infected or healthy as half-edges. Hence, all institutions are of two types and all balls are of four different types. For convenience, we denote them as **S** (solvent), **D** (defaulted) bins, and further $\mathbf{H}^+$ (healthy in), $\mathbf{H}^-$ (healthy out), $\mathbf{I}^+$ (infected in) and $\mathbf{I}^-$ (infected out) balls, respectively.

We start from the set of fundamental defaults $\mathcal{D}_0$, which gives the set of initially defaulted bins and infected balls. At each step, we first remove a uniformly chosen ball of type $\mathbf{I}^-$ and then a uniformly chosen ball from $\mathbf{H}^+ \cup \mathbf{I}^+$. In this process **S** bins may change to **D** bins and, consequently, $\mathbb{H}$ balls may change to **I** balls. We continue the process until there is no more $\mathbf{I}^-$ balls. We now change the description a little by introducing colors for the $\mathbf{I}^-$ balls and life for all in balls from $\mathbf{H}^+ \cup \mathbf{I}^+$. We let all $\mathbf{I}^-$ balls are white and all in balls from $\mathbf{H}^+ \cup \mathbf{I}^+$ are initially alive. We begin by recoloring one random $\mathbf{I}^-$ ball red. Subsequently, in each removal step, we first kill a random in ball from $\mathbf{H}^+ \cup \mathbf{I}^+$ and at the same moment we also recolor a random white ball red. This is repeated until no more white $\mathbf{I}^-$ balls remain.

We next run the above death process in continuous time. We assume that each ball from $\mathbf{H}^+ \cup \mathbf{I}^+$ has an exponentially distributed random lifetime with mean one, independent of all other balls. Namely, if there are ℓ alive in balls remaining, then we wait an exponential time with mean $1/\ell$ until the next pair of interactions. We stop when we should recolor a white ball red but there is no such ball. Let us denote by $W_n(t)$ the number of white $\mathbf{I}^-$ balls at time t. Hence, the above death process ends at the stopping time $\tau_n^\star$ which is the first time when we need to recolor a white ball but there are no white balls left. However, we pretend that we recolor a (nonexistent) white ball at time $\tau_n^\star$ and write $W_n(\tau_n^\star) = -1$. The pseudocode for this default cascade death process is provided in Algorithm 1.

Algorithm 1: Default cascade death process

---

1. **Initialize:**
   (a) Set up the set of fundamentally defaulted institutions $\mathcal{D}_0(\boldsymbol{\epsilon})$.
   (b) Assign the threshold θ_i to each institution $i \in [n]$, according to the distribution $q_{x_i}(.)$.
   (c) Mark all outgoing half-edges originating from $\mathcal{D}_0(\boldsymbol{\epsilon})$ in white.
   (d) Allocate i.i.d. exponential lifetimes with a mean of one to all incoming half-edges.
2. **while** there exists a white outgoing half-edge in the system **do**
   (a) Wait for the next incoming half-edge death (uniformly distributed among all alive incoming half-edges) and remove this half-edge from the set of alive incoming half-edges.
   (b) If this incoming half-edge is connected to institution $i \in [n]$ with threshold θ_i and this is the θ_i-th incoming loss to this institution, then add this institution to the set of defaulted institutions and color all outgoing half-edges connected to this institution in white.
   (c) Select a random white outgoing half-edge and color it in red.

   end

We denote by $I_n^+(t)$, $H_n^+(t)$ and $L_n(t)$ the number of alive (in) balls in $\mathbf{I}^+$, $\mathbf{H}^+$ and $\mathbf{H}^+ \cup \mathbf{I}^+$ at time t, respectively. For $x \in \mathcal{X}, \theta \in \mathbb{N}, \ell = 0, \dots, \theta - 1$, we let $S_{x,\theta,\ell}^{(n)}(t)$ denote the number of solvent institutions (bins) with type x, threshold θ and ℓ defaulted neighbors at time t. Further, let $S_n(t)$ and $D_n(t)$ be the numbers of $\mathbf{S}$ bins and $\mathbf{D}$ bins at time t. Hence, $S_n(\tau_n^\star)$ denotes the final number of solvent institutions. Further, $D_n(\tau_n^\star) = n - S_n(\tau_n^\star) = |\mathcal{D}_{n-1}|$ is the final number of defaulted institutions.

2.3 Limit Theorems

In this section we consider the above dynamic default contagion model (which is now transferred to a death process problem represented by balls-and-bins) and state our main results regarding the limit theorems in the random financial network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$.

We first define some functions that will be used later. Let for $z \in [0,1]$:

$$b(d,z,\ell) := \mathbb{P}(\mathsf{Bin}(d,z)=\ell) = \binom{d}{\ell} z^{\ell}(1-z)^{d-\ell}, \tag{2.2}$$

$$\beta(d,z,\ell) := \mathbb{P}(\mathsf{Bin}(d,z)\geqslant \ell) = \sum_{r=\ell}^{d} \binom{d}{r} z^{r}(1-z)^{d-r}, \tag{2.3}$$

and $\mathsf{Bin}(d,z)$ denotes the binomial distribution with parameters d and z.

2.3.1 Asymptotic Magnitude of Default Contagion

We consider the random financial network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ and assume that the average degrees converges to a finite limit.

Assumption 2.3a. *We assume that, as $n \to \infty$, the average degrees converges and is finite:*

$$\lambda^{(n)} := \sum_{x\in\mathcal{X}} d_x^+ \mu_x^{(n)} = \sum_{x\in\mathcal{X}} d_x^- \mu_x^{(n)} \longrightarrow \lambda := \sum_{x\in\mathcal{X}} d_x^+ \mu_x \in (0,\infty).$$

Note that the finite average degree condition is satisfied for most real-world scale-free financial networks with a shape parameter larger than 2.

For $z \in [0,1]$, we define the functions:

$$f_S(z) := \sum_{x\in\mathcal{X}} \mu_x \sum_{\theta=1}^{d_x^+} q_x(\theta)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big), \quad f_D(z) = 1 - f_S(z),$$

$$f_{H^+}(z) := \sum_{x\in\mathcal{X}} \mu_x \sum_{\theta=1}^{d_x^+} q_x(\theta) \sum_{\ell=d_x^+-\theta+1}^{d_x^+} \ell b\big(d_x^+, z, \ell\big), \quad f_{I^+}(z) = \lambda z - f_{H^+}(z),$$

$$f_W(z) := \lambda z - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=1}^{d_x^+} q_x(\theta)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big).$$

The following theorem states the law of large numbers for the number of solvent banks, defaulted banks, healthy links, infected links and the total number of existing white balls (remaining interactions yielding at least one default) at any time t in the economy $\mathcal{E}_n$ satisfying above regularity assumptions.

Theorem 2.1. *Suppose that Assumptions 2.1-2.3a hold. Let $\tau_n \leqslant \tau_n^\star$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Then for all $x \in \mathcal{X}, \theta = 1, \dots, d_x^+$ and $\ell = 0, \dots, \theta - 1$, we have*

$$\sup_{t\leqslant\tau_n} \Big| \frac{S_{x,\theta,\ell}^{(n)}(t)}{n} - \mu_x q_x(\theta) b\left(d_x^+, 1-e^{-t}, \ell\right) \Big| \xrightarrow{p} 0 \quad as \quad n \to \infty.$$

Further, as $n \to \infty$,

$$\sup_{t \leqslant \tau_n} \Big| \frac{S_n(t)}{n} - f_S(e^{-t}) \Big| \xrightarrow{p} 0, \quad \sup_{t \leqslant \tau_n} \Big| \frac{D_n(t)}{n} - f_D(e^{-t}) \Big| \xrightarrow{p} 0,$$
$$\sup_{t \leqslant \tau_n} \Big| \frac{H_n^+(t)}{n} - f_{H^+}(e^{-t}) \Big| \xrightarrow{p} 0, \quad \sup_{t \leqslant \tau_n} \Big| \frac{I_n^+(t)}{n} - f_{I^+}(e^{-t}) \Big| \xrightarrow{p} 0,$$

and the number of white balls satisfies

$$\sup_{t \leqslant \tau_n} \Big| \frac{W_n(t)}{n} - f_W(e^{-t}) \Big| \xrightarrow{p} 0.$$

Proof. see Section 2.7.1. □

We consider now the stopping time $\tau_n^\star$ which is the first time such that $W_n(\tau_n^\star) = -1$ (becomes negative). Let us define

$$z^\star := \sup\big\{z \in [0,1] : f_W(z) = 0\big\}.$$

We have the following lemma.

Lemma 2.2. *Suppose that Assumptions 2.1-2.3a hold. We have (as $n \to \infty$):*

(i) If $z^\star = 0$ then $\tau_n^\star \xrightarrow{p} \infty$.

(ii) If $z^\star \in (0,1]$ and $z^\star$ is a stable solution, i.e. $f'_W(z^\star) > 0$, then $\tau_n^\star \xrightarrow{p} -\ln z^\star$.

Proof. See Appendix 2.8.2. □

Remark 2.3. *The stable solution of $f_W(t)$ guarantees that the process W_n becomes negative when n is large enough, by Theorem 2.1. If the solution is not stable, W_n reaches some position close to 0, but may not be negative. Then we can not guarantee that the default contagion stops.*

As a corollary of Theorem 2.1 and Lemma 2.2, we next provide the law of large numbers for the final state of default contagion.

Theorem 2.4. *Suppose that Assumptions 2.1-2.3a hold. The final fraction of defaults satisfies:*

(i) If $z^\star = 0$ then asymptotically almost all institutions default during the cascade and $\big|\mathcal{D}_{n-1}\big| = n - o_p(n)$.

(ii) If $z^\star \in (0,1]$ and $z^\star$ is a stable solution, i.e. $f'_W(z^\star) > 0$, then $\frac{|\mathcal{D}_{n-1}|}{n} \xrightarrow{p} f_D(z^\star)$. Further, in this case, for all $x \in \mathcal{X}, \theta = 1, \ldots, d_x^+$ and $\ell = 0, \ldots, \theta - 1$, the final fraction of solvent institutions with type x, threshold θ and ℓ defaulted neighbors satisfies

$$\frac{S_{x,\theta,\ell}^{(n)}}{n} \xrightarrow{p} \mu_x q_x(\theta) b\left(d_x^+, 1 - z^\star, \ell\right).$$

Proof. See Section 2.7.2. □

The above theorem (in a simpler setup) has been used in [20] to provide a resilience condition for contagion in random financial networks. With the notation above, starting from a small fraction ϵ of institutions representing the fundamental defaults, i.e., $\sum_{x\in\mathcal{X}} \mu_x q_x(0) = \epsilon$, the financial network is said to be resilient if $\lim_{\epsilon\to 0} z^\star = 1$; this condition implies that the final fraction of defaults is (w.h.p.) negligible and $|\mathcal{D}_{n-1}| = o_p(n)$. We refer to [12, 20] for the resilience conditions.

The law of large numbers results can be used for quantifying systemic risk in different networks. This can be reflected by various wealth aggregation functions providing indicators for the health of financial systems in different stress scenarios.

2.3.2 Asymptotic Normality of Default Contagion

In order to study the central limit theorems, we need to restrict our attention to the sparse networks regime. Namely, we consider the random financial network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ and assume that degrees sequences satisfy the following moment condition.

Assumption 2.3b. *We assume that for every constant $A > 1$, we have*

$$\sum_{i=1}^{n} A^{d_i^+} = n \sum_{x\in\mathcal{X}} \mu_x^{(n)} A^{d_x^+} = O(n) \quad \text{and} \quad \sum_{i=1}^{n} A^{d_i^-} = n \sum_{x\in\mathcal{X}} \mu_x^{(n)} A^{d_x^-} = O(n).$$

Compared to Assumption 2.3a, this assumption restricts the networks to a sparse regime (e.g., a Core-Periphery financial network or an Erdös-Rényi random graph with finite average degree).

Remark 2.5. *Let (D_n^+, D_n^-) be random variables with joint distribution*

$$\mathbb{P}(D_n^+ = d^+, D_n^- = d^-) = \sum_{x\in\mathcal{X}} \mu_x^{(n)} \mathbb{1}\{d_x^+ = d^+, d_x^- = d^-\},$$

which is the joint distribution of in- and out- degrees for a random node in $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$. Let also (D^+, D^-) be random variables (over nonnegative integers) with joint distribution

$$\mathbb{P}(D^+ = d^+, D^- = d^-) = \sum_{x\in\mathcal{X}} \mu_x \mathbb{1}\{d_x^+ = d^+, d_x^- = d^-\}.$$

Assumption 2.3b can be rewritten as $\mathbb{E}[A^{D_n^+}] = O(1)$ and $\mathbb{E}[A^{D_n^-}] = O(1)$ for each $A > 1$, which in particular implies the uniform integrability of D_n^+ and D_n^-, so

$$\lambda^{(n)} := \sum_{x\in\mathcal{X}} d_x^+ \mu_x^{(n)} = \mathbb{E}[D_n^+] \longrightarrow \mathbb{E}[D^+] = \lambda \in (0, \infty).$$

Similarly, all higher moments converge.

By the construction of the balls-and-bins model, the independency exists between any two different types $x_1 \neq x_2$, at any time t before the final state is reached. Hence we study the asymptotic normality type by type. We first show the following joint convergence theorem for all $x \in \mathcal{X}, \theta = 1, \ldots, d_x^+, \ell = 0, \ldots, \theta - 1$.

Theorem 2.6. *Suppose that Assumptions 2.1-2.3b hold. Let $\tau_n \leqslant \tau_n^\star$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. For all $x \in \mathcal{X}, \theta = 1, \ldots, d_x^+, \ell = 0, \ldots, \theta - 1$ and jointly in $\mathcal{D}[0, \infty)$,*

$$n^{-1/2}\left(S_{x,\theta,\ell}^{(n)}(t \wedge \tau_n) - n\mu_x^{(n)} q_x^{(n)}(\theta) b\big(d_x^+, 1 - e^{-(t\wedge\tau_n)}, \ell\big)\right) \xrightarrow{d} \mathcal{Z}_{x,\theta,\ell}(t \wedge t_0),$$

where $\mathcal{Z}_{x,\theta,\ell}(t)$ is a Gaussian process with mean 0 and variance $\sigma_{x,\theta,\ell}(t)$ given by (2.12).

Proof. See Section 2.7.3, where we also provide the covariance between $\mathcal{Z}^*_{x_1,\theta_1,\ell_1}$ and $\mathcal{Z}^*_{x_2,\theta_2,\ell_2}$, for any two triplets (x_1, θ_1, ℓ_1) and (x_2, θ_2, ℓ_2); (see (2.11)). □

The process $S_{x,\theta,\ell}^{(n)}$ is an elementary process in the network. Other processes can be regarded as aggregated processes of $S_{x,\theta,\ell}^{(n)}$ over different (x, θ, ℓ). With the asymptotic normality of $S_{x,\theta,\ell}^{(n)}$, the asymptotic normality of other processes can be obtained.

In the following theorem, we show the joint asymptotic normality between the total number of solvent institutions, number of defaulted institutions, number of infected and healthy links, and the total number of white balls (controlling the default contagion stopping time) at any time t before the end of default cascade. For $z \in [0, 1]$, we define the functions:

$$\begin{aligned}
f_S^{(n)}(z) &:= \sum_{x\in\mathcal{X}} \mu_x^{(n)} \sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big), \quad f_D^{(n)}(z) = 1 - f_S^{(n)}(z),\\
f_{H^+}^{(n)}(z) &:= \sum_{x\in\mathcal{X}} \mu_x^{(n)} \sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta) \sum_{\ell=d_x^+-\theta+1}^{d_x^+} \ell b\big(d_x^+, z, \ell\big), \quad f_{I^+}^{(n)}(z) = \lambda z - f_{H^+}^{(n)}(z),\\
f_W^{(n)}(z) &:= \lambda^{(n)} z - \sum_{x\in\mathcal{X}} \mu_x^{(n)} d_x^- \sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big).
\end{aligned}$$

For convenience, we set

$$\widehat{f}_\clubsuit^{(n)}(t) = f_\clubsuit^{(n)}(e^{-t}), \text{ for } \clubsuit \in \{S, D, H^+, I^+, W\}.$$

Theorem 2.7. *Suppose that Assumptions 2.1-2.3b hold. Let $\tau_n \leqslant \tau_n^\star$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Then jointly in $\mathcal{D}[0, \infty)$, as $n \to \infty$,*

$$n^{-1/2}\left(\clubsuit_n(t \wedge \tau_n) - n\widehat{f}_\clubsuit^{(n)}(t \wedge \tau_n)\right) \xrightarrow{d} \mathcal{Z}_\clubsuit(t \wedge t_0) \tag{2.4}$$

for $\clubsuit \in \{S, D, H^+, I^+, W\}$, where $\{Z_\clubsuit\}$ are continuous Gaussian processes on $[0, t_0]$ with mean 0 and covariances that satisfy, for $0 \leqslant t \leqslant t_0$ and $\clubsuit, \spadesuit \in \{S, D, H^+, I^+, W\}$,

$$\operatorname{Cov}\big(\mathcal{Z}_\clubsuit(t), \mathcal{Z}_\spadesuit(t)\big) = \sigma_{\clubsuit,\spadesuit}(e^{-t}),$$

where the form of $\sigma_{\clubsuit,\spadesuit}(x)$ are given by (2.56)-(2.60) in Appendix 2.8.7.

Proof. See Section 2.7.4. Note that since $D_n(t) = n - S_n(t)$, it is easy to transfer the result to D by setting $\sigma_{D,D} = \sigma_{S,S}$ and $\sigma_{D,\clubsuit} = -\sigma_{S,\clubsuit}$. Further, since $I_n(t) = L_n(t) - H_n^+(t)$, computing the covariances of I^+ is similar to H^+ and is omitted from the proof. We only provide the covariances for $\clubsuit, \spadesuit \in \{S, H^+, W\}$. ☐

We have obtained asymptotic normality for all processes of interest. Note that the covariance functions between them are obtained in explicit forms, using the observable features of the network. This allows us to test the approximation performance in networks with size n, when n is not so large.

Our central limit theorems can be used to provide a confidence interval for the fraction of defaults in finite financial networks. Figure 2.1 displays the 95% confidence interval for the fraction of defaults in a 6-regular financial network ($d^+ = d^- = 6$) (plotted against $z = e^{-t}$) during the default cascade process. As expected, when the network size is larger, the interval size becomes smaller.

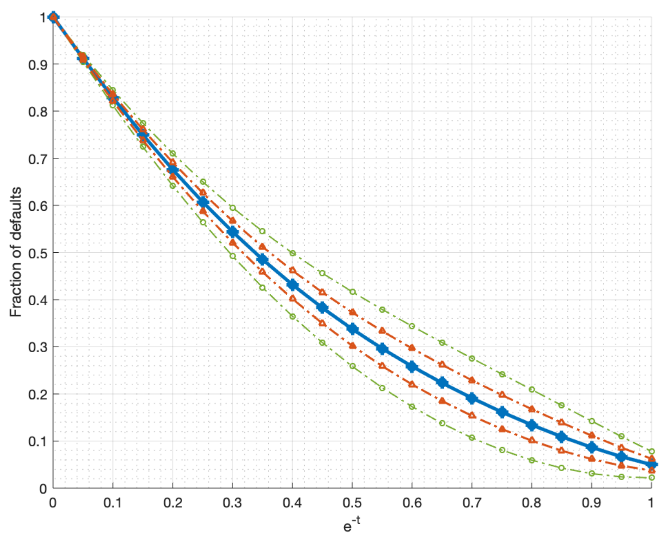


Figure 2.1: 95% confidence interval of the fraction of defaulted institutions. The blue solid line is the limit, the green dash line is the bounds for network size $n = 300$, and the red dash line is the bounds for network size $n = 1500$. Here, $d^+ = d^= = 6$ and the threshold distribution is $q(0) = 0.05, q(1) = 0.05, q(2) = 0.1, q(3) = 0.1, q(4) = 0.15, q(5) = 0.25$ and $q(6) = 0.3$.

Let us define

$$s_{x,\theta,\ell}(z) := \mu_x q_x(\theta) b\big(d_x^+, 1-z, \ell\big), \quad s_{x,\theta,\ell}^{(n)}(z) := \mu_x^{(n)} q_x^{(n)}(\theta) b\big(d_x^+, 1-z, \ell\big),$$

and let $\widehat{z}_n$ be the largest $z \in [0,1]$ such that $f_W^{(n)}(z) = 0$. As a corollary of Theorem 2.7 and Lemma 2.2, we have the following result regarding the final state of default contagion.

Theorem 2.8. *Suppose that Assumptions 2.1-2.3b hold. Let* $t^\star = -\ln z^\star$*,If* $z^\star \in (0,1]$ *and* $z^\star$ *is a stable solution, i.e.* $\alpha := f_W'(z^\star) > 0$*, then we have*

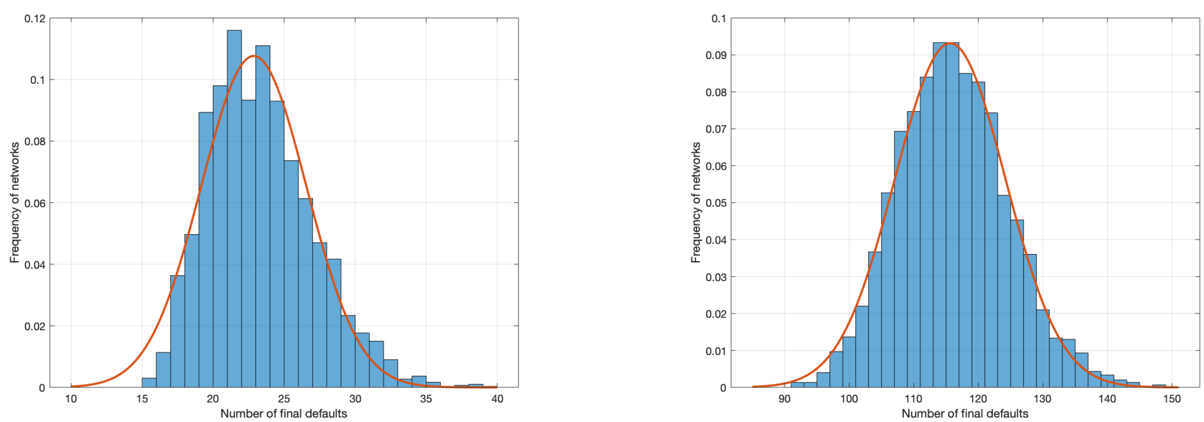
$$n^{-1/2}(\clubsuit_n(\tau_n^\star) - n f_\clubsuit^{(n)}(\widehat{z}_n)) \xrightarrow{d} \mathcal{Z}_\clubsuit(t^\star) - \alpha^{-1} f_\clubsuit'(z^\star)\mathcal{Z}_W(t^\star), \tag{2.5}$$

for $\clubsuit \in \{S, D, H^+, I^+, W\}$*, where the limit distributions compose a Gaussian vector. Furthermore,* $\widehat{z}_n \to z^\star$ *and, for all* $x \in \mathcal{X}$*,* $0 \leqslant \ell < \theta \leqslant d_x^+$*,*

$$n^{-1/2}(S_{x,\theta,\ell}^{(n)}(\tau_n^\star) - n s_{x,\theta,\ell}^{(n)}(\widehat{z}_n)) \xrightarrow{d} \mathcal{Z}_{x,\theta,\ell}^*(t^\star) - \alpha^{-1} s_{x,\theta,\ell}'(z^\star)\mathcal{Z}_W(t^\star). \tag{2.6}$$

Proof. See Section 2.7.5. □

To study the convergence of our central limit theorems numerically, we consider in Figure 2.2 networks with finite size n and simulate the final fraction of defaulted institutions by using a Monte-Carlo method. To see how the distributions of final fraction of defaults come close to the Gaussian distributions as n becomes large, we run 3000 times the default cascade process of Figure 2.1 in different 6-regular networks chosen uniformly at random among all 6-regular (directed) networks. We count how many institutions default at the end of each simulation and then produce the histograms. Figure 2.2 displays the obtained histograms with two different network sizes $n = 300$ and $n = 1500$.



(a) Network size n=300

(b) Network size n=1500

Figure 2.2: Histograms of 3000 times Monte-Carlo simulation for the number of final defaults in regular financial networks with size $n = 300, 1500$.

2.4 Quantifying Systemic Risk

In order to determine the health of the financial network, we consider now a *systemic risk measure* applied to the (random) financial network, introduced in previous sections. These measures are decomposed as $\rho \circ \Gamma$ for a stand-alone risk measure (usually assumed convex) ρ and an aggregation function $\Gamma = \Gamma(\boldsymbol{\epsilon})$ for losses under the stress scenario $\boldsymbol{\epsilon}$. This was first introduced in [87, 159]; see also [21, 118]. The following three aggregation functions have been considered in the literature. At time t, for the economy $\mathcal{E}_n$ and given shock scenario $\boldsymbol{\epsilon}$, we let:

- **Number of solvent banks**: $\Gamma_n^{\#}(t) := S_n(t) = n - D_n(t)$.
- **External wealth**: Let $\bar{\Gamma}_n^{\odot}$ denote the total external wealth to society if there is no default in the financial system (small shock regime). We define the external wealth (societal) aggregation function as $\Gamma_n^{\odot}(t) := \bar{\Gamma}_n^{\odot} - \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} D_x^{(n)}(t)$, where $D_x^{(n)}(t) = n\mu_x^{(n)} - \sum_\theta \sum_{\ell=0}^{\theta-1} S_{x,\theta,\ell}^{(n)}(t)$ denotes the total number of defaulted institutions with type $x \in \mathcal{X}$ at time t. Note that (for simplicity) we assume a bounded constant type-dependent societal loss $\bar{L}_x^{\odot}$ over each defaulted institution.
- **System-wide wealth** : Let $\bar{\Gamma}_n^{\Diamond}$ denote the total wealth in the financial system if there is no default in the system. We define the system-wide aggregation function as

$$\Gamma_n^{\Diamond}(t) := \bar{\Gamma}_n^{\Diamond} - \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} D_x^{(n)}(t) - \sum_{x\in\mathcal{X}} \bar{L}_x^{\Diamond} \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell S_{x,\theta,\ell}^{(n)}(t).$$

  For each type $x \in \mathcal{X}$, we consider a bounded fixed (type-dependent) societal cost $\bar{L}_x^{\odot}$ for defaulted institutions and a bounded fixed (host institutions' type-dependent) cost $\bar{L}_x^{\Diamond}$ over each defaulted links.

For the aggregation function $\Gamma_n^{\#}(t)$, we already stated the limit theorems in Section 2.3. Since the societal aggregation function $\Gamma_n^{\odot}$ can be seen as a particular case of system-wide aggregation function $\Gamma_n^{\Diamond}$ (by setting $\bar{L}_x^{\Diamond} = 0$), we only state limit theorems for $\Gamma_n^{\Diamond}$.

To this purpose, it is natural to assume that $\bar{\Gamma}_n^{\Diamond}/n \to \bar{\Gamma}^{\Diamond}$ when the size of network $n \to \infty$. Let us define

$$f_{\Diamond}^{(n)}(z) := \bar{\Gamma}_n^{\Diamond}/n - \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} f_D^{(n)}(z) - \sum_{x\in\mathcal{X}} \bar{L}_x^{\Diamond} \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell s_{x,\theta,\ell}^{(n)}(z),$$

$$f_{\Diamond}(z) := \bar{\Gamma}^{\Diamond} - \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} f_D(z) - \sum_{x\in\mathcal{X}} \bar{L}_x^{\Diamond} \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell s_{x,\theta,\ell}(z).$$

Similarly we also set

$$\widehat{f}_{\Diamond}^{(n)}(t) = f_{\Diamond}^{(n)}(e^{-t}), \qquad \widehat{f}_{\Diamond}(t) = f_{\Diamond}(e^{-t}).$$

We next consider the central limit theorems for the societal and system-wide aggregation functions. By applying Theorem 2.1 and Theorem 2.4, the following holds.

Theorem 2.9. *Suppose that Assumptions 2.1-2.3a hold. Let $\tau_n \leqslant \tau_n^\star$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Then, as $n \to \infty$,*

$$\sup_{t\leqslant\tau_n} \Big| \frac{\Gamma_n^{\Diamond}(t)}{n} - f_{\Diamond}(e^{-t}) \Big| \xrightarrow{p} 0, \tag{2.7}$$

and under Assumption 2.3b, jointly in $\mathcal{D}[0,\infty)$,

$$n^{-1/2}\left(\Gamma_n^{\Diamond}(t\wedge\tau_n) - n\widehat{f}_{\Diamond}^{(n)}(t\wedge\tau_n)\right) \xrightarrow{d} \mathcal{Z}_{\Diamond}(t\wedge t_0),$$

where $\mathcal{Z}_{\Diamond}$ *is a continuous Gaussian process on* $[0,t_0]$ *with mean 0 and variance* $\sigma_{\Diamond}(t)$ *given by* (2.61) *in Appendix 2.8.8.*

Moreover, the final (system-wide) aggregation functions satisfy (under Assumptions 2.1-2.3a):

(i) If $z^\star = 0$ *then asymptotically almost all institutions default during the cascade and*

$$\frac{\Gamma_n^{\Diamond}(\tau_n^\star)}{n} \xrightarrow{p} \bar{\Gamma}^{\Diamond} - \sum_{x\in\mathcal{X}} \mu_x \bar{L}_x^{\odot}.$$

(ii) If $z^\star \in (0,1]$ *and* $z^\star$ *is a stable solution, i.e.* $f_W'(z^\star) > 0$, *then* $\frac{\Gamma_n^{\Diamond}(\tau_n^\star)}{n} \xrightarrow{p} f_{\Diamond}(z^\star)$ *and, under Assumption 2.3b,*

$$n^{-1/2}\left(\Gamma_n^{\Diamond}(\tau_n^\star) - nf_{\Diamond}^{(n)}(\widehat{z}_n)\right) \xrightarrow{d} \mathcal{Z}_{\Diamond}^\star,$$

where $\mathcal{Z}_{\Diamond}^\star$ *is a centered Gaussian random variable with variance* $\sigma_{\Diamond}^\star$ *given by* (2.62).

Proof. See Section 2.7.6. □

2.4.1 Numerical analysis on heterogeneity and stability

In this section, we investigate the impact of network structure heterogeneity on the final size of the default cascade. To simplify the analysis, we assume that the out-degree is equal to the in-degree for all nodes. We consider the following three types of networks:

- **Regular networks**: These exhibit regular, symmetric linkages among nodes. All nodes, considered identical for our study, create a homogeneous, symmetric network.
- **Erdös-Rényi networks**: In this model, every pair of nodes (a potential directed link) independently forms a connection with a fixed probability $p_n \in (0,1)$, such that $np_n \to \lambda$ as $n\to\infty$. This differs from regular networks as each node pair has a potential edge based on a specific probability, infusing heterogeneity into the degree distribution. In particular, the degree distribution converges to a Poisson distribution with parameter λ, with the asymptotic probability mass function of degree given by $\mathbb{P}(D=k) = e^{-\lambda}\lambda^k/k!$.
- **Scale-free networks**: These are prevalent in many real-world financial network systems. Scale-free networks possess a degree distribution following a power law. This is expressed as $\mathbb{P}(D = k) \sim ck^{-\eta}$, where $c>0$ is a normalizing constant and $\eta > 1$ is a control parameter.

We set the parameters $\lambda = 5$ and $\eta = 1.2$. To reduce the complexity of the simulation, we assume that the degrees are upper-bounded by $d_{\max} = 23$. These parameter choices result in both the scale-free and Erdös-Rényi networks having an average degree very close to 5. In Figure 2.3, we compare the final fraction of defaults in these networks to a regular network with a degree of 5. We assess the final defaults in these three distinct networks under different initial shocks, measured as a percentage of asset loss for each agent. Our numerical framework employs i.i.d. distributed Pareto losses.

From Figure 2.3, we note that for small shocks (less than 0.2), the performance of the three networks is quite similar. However, as the shock size increases, the scale-free network is the first to leap to a larger default fraction, indicating it has the smallest critical value for the shock. The ER network follows, and the regular network displays the largest critical value. Interestingly, with larger shocks, the regular network exhibits the highest default fraction among the three, followed by the ER network, while the scale-free network shows the lowest. These observations lead to a conclusion: networks with low heterogeneity are more resilient to small shocks, but their resistance to larger shocks increases with heterogeneity.

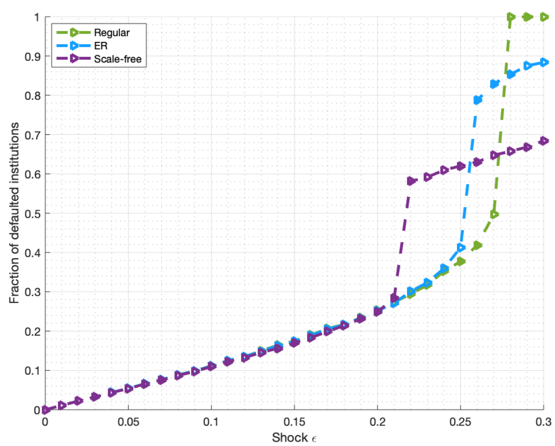


Figure 2.3: Final faction of defaults for regular, Erdös-Rényi (ER) and scale-free networks.

2.5 Targeting Interventions in Financial Networks

In this section we consider a planner (lender of last resort or government) who seeks to minimize the systemic risk at the beginning of the financial contagion, after an exogenous macroeconomic shock $\boldsymbol{\epsilon}$, subject to a budget constraint. As discussed in Section 2.4, we assume that the systemic risk is represented by $\rho(\Gamma_n^{\diamondsuit})$, for some convex function ρ applied to the system-wide wealth. Note that since we study the interventions for a given shock $\boldsymbol{\epsilon}$, the uncertainty (in stress scenario) for the risk measure ρ is only on the network structure (which is assumed to be uniformly at random). The planner only has information regarding the type of each institution and, consequently, the institutions' threshold distributions. Hence, the planner' decision is only based on the type of each institution. The timeline is as follows. At time $t = 0$, the financial network is subject to an economic shock $\boldsymbol{\epsilon}$. At time $t = 1$, the planer (observing the external shock $\boldsymbol{\epsilon}$) computes the threshold distribution $q_x(.)$ for each $x \in \mathcal{X}$. Then she makes decisions, under some budget constraint, on the number (fraction) of interventions over all defaulted links leading to any institution with any type $x \in \mathcal{X}$. When the planner intervenes on a defaulting bank, its threshold (distance to default) increases by 1. These interventions will be type-dependent and at random over all defaulted links directing to the same type institutions.

For $x \in \mathcal{X}$, let us denote by $\alpha_x^{(n)}$ the planner intervention decision on the fraction of the saved links directing to any institution of type $x \in \mathcal{X}$. We assume that $\alpha_x^{(n)} \to \alpha_x$ for all $x \in \mathcal{X}$, and some constants α_x independent of n. Let $\boldsymbol{\alpha}_n = \left\{\alpha_x^{(n)}\right\}_{x\in\mathcal{X}}$, and, let $\Gamma_n^{\Diamond}(\boldsymbol{\alpha}_n)$ denote the system-wide wealth under the intervention decision $\boldsymbol{\alpha}_n$. Further, $S_{x,\theta,\ell}^{(n)}(\boldsymbol{\alpha}_n)$ denotes the number of solvent banks with type x, threshold θ and ℓ defaulted neighbors under the intervention decision $\boldsymbol{\alpha}_n$. Similarly, $D_n(\boldsymbol{\alpha}_n)$ denotes the total number of defaults under intervention $\boldsymbol{\alpha}_n$.

Let $C_x \in \mathbb{R}^+$ denote the cost associated to saving any defaulted link leading to an institution of type $x \in \mathcal{X}$. We assume that C_x is a bounded function. We denote by $\Phi_n(\boldsymbol{\alpha}_n)$ the total cost associated to the planner for the intervention strategy $\boldsymbol{\alpha}_n$.

We state below a limit theorem on the number of solvent institutions, defaulted institutions, the total aggregate wealth of the financial system and the total cost of intervention for the planner, under the intervention decision $\boldsymbol{\alpha}_n$. Let us define

$$f_W^{(\boldsymbol{\alpha})}(z) := \lambda z - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=1}^{d_x^+} q_x(\theta)\beta\big(d_x^+, \alpha_x + (1-\alpha_x)z, d_x^+ - \theta + 1\big),$$

and,

$$z_{\boldsymbol{\alpha}}^\star := \sup\big\{z \in [0,1] : f_W^{(\boldsymbol{\alpha})}(z) = 0\big\}. \tag{2.8}$$

Theorem 2.10. *Suppose that Assumptions 2.1-2.3a hold. Let $\boldsymbol{\alpha}_n \to \boldsymbol{\alpha}$ as $n \to \infty$. If $z_{\boldsymbol{\alpha}}^\star$ is a stable solution, then as $n \to \infty$:*

(i) For all $x \in \mathcal{X}, \theta = 1, \dots, d_x^+$ and $\ell = 0, \dots, \theta - 1$, the final fraction of solvent institutions with type x, threshold θ and ℓ defaulted neighbors under intervention $\boldsymbol{\alpha}_n$ converges to

$$\frac{S_{x,\theta,\ell}^{(n)}(\boldsymbol{\alpha}_n)}{n} \xrightarrow{p} s_{x,\theta,\ell}^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star) := \mu_x q_x(\theta) b\left(d_x^+, (1-\alpha_x)(1-z_{\boldsymbol{\alpha}}^\star), \ell\right).$$

(ii) The total number of defaulted institutions under intervention $\boldsymbol{\alpha}_n$ converges to:

$$\frac{D_n(\boldsymbol{\alpha}_n)}{n} \xrightarrow{p} f_D^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star) := 1 - \sum_{x\in\mathcal{X}} \mu_x \sum_{\theta=1}^{d_x^+} q_x(\theta)\beta\big(d_x^+, \alpha_x + (1-\alpha_x)z_{\boldsymbol{\alpha}}^\star, d_x^+ - \theta + 1\big).$$

(iii) The system-wide wealth under the intervention decision $\boldsymbol{\alpha}_n$ converges to

$$\frac{\Gamma_n^{\Diamond}(\boldsymbol{\alpha}_n)}{n} \xrightarrow{p} f_{\Diamond}^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star) := \bar{\Gamma}^{\Diamond} - \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} f_D^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star) - \sum_{x\in\mathcal{X}} \bar{L}_x^{\Diamond} \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell s_{x,\theta,\ell}^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star).$$

(iv) The total cost of interventions $\boldsymbol{\alpha}_n$ for the planner converges to

$$\frac{\Phi_n(\boldsymbol{\alpha}_n)}{n} \xrightarrow{p} \phi(z^\star_{\boldsymbol{\alpha}}) := \sum_{x\in\mathcal{X}} \mu_x \alpha_x C_x \sum_{\ell=1}^{d_x^+} \ell b\left(d_x^+, 1 - z^\star_{\boldsymbol{\alpha}}, \ell\right).$$

Proof. See Section 2.7.7. $\square$

We conclude that, as $n \to \infty$, the planner optimal decision problem simplifies to

$$\max_{\boldsymbol{\alpha}} f_\Diamond^{(\boldsymbol{\alpha})}(z^\star_{\boldsymbol{\alpha}}) := \bar{\Gamma}^\Diamond - \sum_{x\in\mathcal{X}} \bar{L}_x^\odot f_D^{(\boldsymbol{\alpha})}(z^\star_{\boldsymbol{\alpha}}) - \sum_{x\in\mathcal{X}} \bar{L}_x^\Diamond \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell s^{(\boldsymbol{\alpha})}_{x,\theta,\ell}(z^\star_{\boldsymbol{\alpha}}),$$

$$\text{subject to} \quad \phi(z^\star_{\boldsymbol{\alpha}}) := \sum_{x\in\mathcal{X}} \mu_x \alpha_x C_x \sum_{\ell=1}^{d_x^+} \ell b\left(d_x^+, 1 - z^\star_{\boldsymbol{\alpha}}, \ell\right) \leqslant C,$$

for some budget constraint $C > 0$ and $z^\star_{\boldsymbol{\alpha}}$ given by (2.8).

2.6 Default cascade process

In this section, we study the default cascade processes in the financial network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$.

We consider a nonnegative matrix (ν_{ij}) which represents the frequency of meetings between any connected (ordered) pair of agents $i, j \in [n]$. Assume:

- Each pair of agents i, j meet (and interact) at the jump times of a Poisson process of rate $\nu_{ij} > 0$, independently from all other meetings in the network.
- Each pair of agents i, j with rate $\nu_{ij} = 0$ never meet, which means that agents i, j are not directly connected to each other (there is no liability from i towards j).

This collection of Poisson processes are called the meeting process and the matrix (ν_{ij}) specifies the meeting model.

We now introduce the default cascade process. Starting from the set of fundamentally insolvent agents $\mathbf{D}(0) = \{i \in [n] : c_i < 0\}$ and initial capitals $\boldsymbol{C}(0) = (c_1, \dots, c_n)$, each pair of (connected) agents in the financial network, interacts at random times associated to the above meeting process and update their states (solvency/default state or interbank liabilities). If, at the meeting time, the debtor agent is solvent, the two agents continue to interact and update their states. In this case, the capitals of these two agents are assumed to be unchanged. Otherwise, when a defaulted agent meets its creditor agent, the creditor agent receives a random loss with distribution depending on its characteristics and the the two agents stop meeting each other. Namely, if agent i is defaulted at the meeting time with agent j (where $i \to j$), then it brings a random amount of loss, denoted by L_{ij}, to j. In this case, the capital of agent j is reduced by the loss L_{ij}. We assume that the losses that agent j receives from its debtors are i.i.d positive bounded, depending only on the characteristics of agent j.

We write $\boldsymbol{C}(t) = (c_1(t), \ldots, c_n(t))$ for the time-t configuration of capitals in the above default cascade process, with $c_i(0) = c_i$ for all $i \in [n]$. Then, it is easy to show that $\boldsymbol{C}(t)$ is a continuous-time Markov process which at some almost surely (a.s.) finite random time $\tau_n^\star$, reaches some absorbing configuration $\boldsymbol{C}^\star$ in which there is some random set of agents who are solvent (i.e., with positive capital).

Under the same assumptions on the type distribution and default threshold distribution as in previous sections, we set the following second moment condition on the degree sequence.

Assumption 2.4. *We assume that (as $n \to \infty$) $\sum_{i\in[n]}(d_i^+ + d_i^-)^2 = O(n)$.*

We consider the default cascade process when the meeting times are i.i.d. exponential random variables with parameter (normalized to) one. Namely, we assume that

$$\nu_{ij} = \mathbb{1}\{(i,j) \text{ is an edge of } \mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)\}.$$

The main result is based on studying the time-change default cascade process and using some limit theorems from previous sections on this time-changed Markov process. The result is as following.

Theorem 2.11. *Consider the default cascade process in the random network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ satisfying Assumption 2.4, when the meeting times are i.i.d. exponential random variables with parameter one. If $z^\star$ is a stable solution, then we have the following:*

(i) There exists a unique continuously differentiable function $v : [0,\infty) \to (z^\star, 1]$ satisfying the differential equation

$$\frac{d}{dt}v(t) = -\frac{f_W(v(t))}{\lambda}, \qquad v(0) = 1, \tag{ODE}$$

and, $v(t) \searrow z^\star$ as $t \to \infty$.

(ii) For all $x \in \mathcal{X}, \theta = 1, \ldots, d_x^+$ and $\ell = 0, \ldots, \theta - 1$, as $n \to \infty$,

$$\sup_{t\geqslant 0}\Big|\frac{S_{x,\theta,\ell}^{(n)}(t)}{n} - \mu_x q_x(\theta) b\left(d_x^+, 1 - v(t), \ell\right)\Big| \xrightarrow{p} 0.$$

Further, as $n \to \infty$,

$$\sup_{t\geqslant 0}\Big|\frac{S_n(t)}{n} - f_S(v(t))\Big| \xrightarrow{p} 0 \quad \text{and} \quad \sup_{t\geqslant 0}\Big|\frac{D_n(t)}{n} - f_D(v(t))\Big| \xrightarrow{p} 0.$$

Let $T_n(\ell)$ be the (random) time that the random financial network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ faces its ℓ-th loss (ℓ-th meeting time containing at least one default). Let also $L_n^\star$ denote the final number of losses in network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$. Then we have the following theorem.

Theorem 2.12. *Consider the default cascade process in the random network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ satisfying Assumption 2.4, when the meeting times are i.i.d. exponential random variables with parameter one. Then, as $n \to \infty$, $L_n^\star/n \xrightarrow{p} \lambda(1 - z^\star)$, and, for all $0 \leqslant a < b < 1 - z^\star$,*

$$T_n(b\lambda^{(n)}n) - T_n(a\lambda^{(n)}n) \xrightarrow{p} \int_{1-b}^{1-a} \frac{\lambda}{f_W(x)}dx.$$

The proof of above theorems are provided in §2.7.8.

2.7 Proof of Main Theorems

This section contains the proofs of all the theorems in previous sections. We provide in Appendix 2.8.1 proofs of lemmas and recall some useful preliminary results on death processes and martingale theory.

2.7.1 Proof of Theorem 2.1

We denote by $U^{(n)}_{x,\theta,s}(t)$ the number of bins (institutions) with type $x \in \mathcal{X}$, threshold θ and s alive (in-) balls at time t. Let $N^{(n)}_{x,\theta}$ denote the (random) number of bins with type x and threshold θ. Let also $N^{(n)}_{x} = \sum_{\theta} N^{(n)}_{x,\theta}$ denote the number of bins with type x. We first state the following lemma on the convergence of $U^{(n)}_{x,\theta,s}(t)$.

Lemma 2.13. *Let $\tau_n \leqslant \tau_n^\star$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Under Assumption 2.3a, for all $x \in \mathcal{X}, \theta = 1, \dots, d_x^+$ and $\ell = 0, \dots, \theta - 1$, we have (as $n \to \infty$)*

$$\sup_{t \leqslant \tau_n} \Big| \frac{U^{(n)}_{x,\theta,\ell}(t)}{n} - \mu_x q_x(\theta) b\left(d_x^+, e^{-t}, \ell\right) \Big| \xrightarrow{p} 0. \tag{2.9}$$

Further,

$$\sup_{t \leqslant \tau_n} \sum_{x \in \mathcal{X}} (d_x^+ + d_x^-) \sum_{\theta=1}^{d_x^+} \sum_{s=d_x^+-\theta+1}^{d_x^+} \left| U^{(n)}_{x,\theta,s}(t)/n - \mu_x q_x(\theta) b(d_x^+, e^{-t}, s) \right| \xrightarrow{p} 0. \tag{2.10}$$

Proof. The proof is based on the death process Lemma 2.19 and provided in Section 2.8.3. $\square$

We now continue with the proof of Theorem 2.1.

Consider $S^{(n)}_{x,\theta,\ell}$, the number of solvent institutions with type x, threshold θ and $\ell = 0, \dots, \theta - 1$ defaulted neighbors at time t. By definition, $S^{(n)}_{x,\theta,\ell}(t) = U^{(n)}_{x,\theta,d_x^+-\ell}(t)$. Hence, by (2.9), we obtain that

$$\sup_{t \leqslant \tau_n} \Big| \frac{S^{(n)}_{x,\theta,\ell}(t)}{n} - \mu_x q_x(\theta) b\left(d_x^+, 1 - e^{-t}, \ell\right) \Big| \xrightarrow{p} 0, \quad \text{as} \quad n \to \infty.$$

The total number of solvent institutions at time t satisfies

$$S_n(t) = \sum_{x \in \mathcal{X}} \sum_{\theta=1}^{d_x^+} \sum_{s=d_x^+-\theta+1}^{d_x^+} U^{(n)}_{x,\theta,s}(t),$$

which is dominated by $\sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+} U_{x,\theta,s}^{(n)}(t)$. Then, by Lemma 2.13 and the convergence result (2.10), we obtain $\sup\limits_{t\leqslant\tau_n}\left|\frac{S_n(t)}{n} - f_S(e^{-t})\right| \xrightarrow{p} 0$.

Further, from $D_n(t) = n - S_n(t)$, the number of defaulted institutions at time t satisfies $\sup\limits_{t\leqslant\tau_n}\left|\frac{D_n(t)}{n} - f_D(e^{-t})\right| \xrightarrow{p} 0$.

Observe also that the total number of healthy in links at time t is given by

$$H_n^+(t) = \sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+} sU_{x,\theta,s}^{(n)}(t),$$

which is also dominated by $\sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+} U_{x,\theta,s}^{(n)}(t)$ and again by Lemma 2.13, we obtain $\sup\limits_{t\leqslant\tau_n}\left|\frac{H_n^+(t)}{n} - f_{H^+}(e^{-t})\right| \xrightarrow{p} 0$.

Moreover, the number of in balls from $\mathbf{I}^+$ at time t satisfies $I_n^+(t) = L_n(t) - H_n^+(t)$. By the construction of the balls-and-bins model, it is easily seen that $L_n(t)$ is a pure death process. It follows by Lemma 2.19 and Assumption 2.3a that $\sup_{t\geqslant 0}\left|L_n(t)/n - \lambda e^{-t}\right| \xrightarrow{p} 0$. We therefore obtain (by definition $f_{I^+}(z) = \lambda z - f_{H^+}(z)$) $\sup\limits_{t\leqslant\tau_n}\left|\frac{I_n^+(t)}{n} - f_{I^+}(e^{-t})\right| \xrightarrow{p} 0$.

Finally, the total number of white (out) balls at time t satisfies $W_n(t) = L_n(t) - H_n^-(t)$, where $H_n^-(t)$ denotes the total number of healthy (out) balls at time t, given by

$$H_n^-(t) = \sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+} d_x^- U_{x,\theta,s}^{(n)}(t).$$

This is again dominated by $\sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+} U_{x,\theta,s}^{(n)}(t)$. Then Lemma 2.13 and Assumption 2.3a imply that the number of white balls satisfies $\sup\limits_{t\leqslant\tau_n}\left|\frac{W_n(t)}{n} - f_W(e^{-t})\right| \xrightarrow{p} 0$. This completes the proof of Theorem 2.1.

2.7.2 Proof of Theorem 2.4

By Theorem 2.1, it follows that $\frac{S_n(\tau_n^\star)}{n} = f_S(e^{-\tau_n^\star}) + o_p(1)$. If $z^\star = 0$ then, by Lemma 2.2, $\tau_n^\star \xrightarrow{p} \infty$. So $e^{-\tau_n^\star} \xrightarrow{p} 0$ and, since $f_S(0) = 0$, it follows by the continuity of f_S that $f_S(e^{-\tau_n^\star}) \xrightarrow{p} 0$. We therefore have $S_n(\tau_n^\star) = o_p(n)$. This implies that $|\mathcal{D}_{n-1}| = n - S_n(\tau_n^\star) = n - o_p(n)$ and, as desired, asymptotically almost all institutions default.

If $z^\star \in (0,1]$ and $f_W'(z^\star) > 0$, then by Lemma 2.2, we have $e^{-\tau_n^\star} \xrightarrow{p} z^\star$. Moreover, the continuity of f_D implies that $f_D(e^{-\tau_n^\star}) \xrightarrow{p} f_D(z^\star)$. Hence, we have by Theorem 2.1 that

$$\frac{|\mathcal{D}_{n-1}|}{n} = \frac{D_n(\tau_n^\star)}{n} \xrightarrow{p} f_D(z^\star).$$

Now using the first statement of Theorem 2.1 and the continuity of $\mu_x q_x(\theta) b(d_x^+, 1-z, \ell)$ on z, we obtain for all $x \in \mathcal{X}, \theta = 1, \ldots, d_x^+$ and $\ell = 0, \ldots, \theta - 1$, the final fraction of solvent institutions with type x, threshold θ and ℓ defaulted neighbors satisfies

$$\frac{S_{x,\theta,\ell}^{(n)}}{n} \xrightarrow{p} \mu_x q_x(\theta) b\left(d_x^+, 1-z^\star, \ell\right).$$

This completes the proof of Theorem 2.4.

2.7.3 Proof of Theorem 2.6

Recall that $U_{x,\theta,s}^{(n)}(t)$ denotes the number of bins (institutions) with type $x \in \mathcal{X}$, threshold θ and s alive (in-) balls at time t. Further, we let $V_{x,\theta,s}^{(n)}(t)$ denote the number of bins (institutions) with type $x \in \mathcal{X}$, threshold θ and at least s alive balls at time t, so that $V_{x,\theta,s}^{(n)}(t) = \sum_{\ell \geqslant s} U_{x,\theta,\ell}^{(n)}(t)$.

We first study the stochastic process $V_{x,\theta,s}^{(n)}$ for a given $x \in \mathcal{X}$ and integers θ, s. For all possible triple (x, θ, s), we define

$$V_{x,\theta,s}^{*(n)}(t) := n^{-1/2}\big(V_{x,\theta,s}^{(n)}(t) - n\mu_x^{(n)} q_x^{(n)}(\theta)\beta(d_x^+, e^{-t}, s)\big),$$

and

$$N_{x,\theta}^{*(n)} := n^{-1/2}\big(N_{x,\theta}^{(n)} - n\mu_x^{(n)} q_x^{(n)}(\theta)\big).$$

We then have the following lemma.

Lemma 2.14. *Let $\tau_n \leqslant \tau_n^\star$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Under Assumption 2.3b, we have that for all couple (x, θ), jointly as $n \to \infty$, $N_{x,\theta}^{*(n)} \xrightarrow{d} \mathcal{Y}_{x,\theta}^*$, where all $\mathcal{Y}_{x,\theta}^*$ are Gaussian random variables with mean 0 and covariances*

$$\mathrm{Cov}(\mathcal{Y}_{x_1,\theta_1}^*, \mathcal{Y}_{x_2,\theta_2}^*) = \psi_{x,\theta_1,\theta_2} \mathbb{1}\{x_1 = x_2\},$$

where $\psi_{x,\theta,\theta} := \mu_x q_x(\theta)(1 - q_x(\theta))$, $\quad \psi_{x,\theta_1,\theta_2} := -\mu_x q_x(\theta_1) q_x(\theta_2) \quad$ for all $\quad \theta_1 \neq \theta_2$.

Further for all triple (x, θ, s), jointly in $\mathcal{D}[0, \infty)$, as $n \to \infty$, $V_{x,\theta,s}^{*(n)}(t \wedge \tau_n) \xrightarrow{d} \mathcal{Z}_{x,\theta,s}^*(t \wedge t_0)$, where all $\mathcal{Z}_{x,\theta,s}^*(t)$ are continuous Gaussian processes with mean 0 and covariances

$$\begin{aligned}
\mathrm{Cov}\big(\mathcal{Z}_{x_1,\theta_1,s_1}^*(t), \mathcal{Z}_{x_2,\theta_2,s_2}^*(t)\big) &= 0, \quad \textit{for all} \quad x_1 \neq x_2, \\
\mathrm{Cov}\big(\mathcal{Z}_{x,\theta_1,s_1}^*(t), \mathcal{Z}_{x,\theta_2,s_2}^*(t)\big) &= \widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(e^{-t}), \quad \textit{for all} \quad \theta_1 \neq \theta_2, \\
\mathrm{Cov}\big(\mathcal{Z}_{x,\theta,s_1}^*(t), \mathcal{Z}_{x,\theta,s_2}^*(t)\big) &= \widehat{\sigma}_{x,\theta,\theta,s_1,s_2}(e^{-t}) + \widetilde{\sigma}_{x,\theta,s_1,s_2}(e^{-t}),
\end{aligned}$$

where $\widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(e^{-t}) := \beta(d_x^+, e^{-t}, s_1)\beta(d_x^+, e^{-t}, s_2)\psi_{x,\theta_1,\theta_2}$, and $\widetilde{\sigma}_{x,\theta,s_1,s_2} = \widetilde{\sigma}_{x,\theta,s_2,s_1}$ with

$$\widetilde{\sigma}_{x,\theta,s,s+k}(y) := \frac{1}{2} y^{2s+k} \sum_{j=s+k}^{d_x^+} \binom{j-1}{s-1}\binom{j-1}{s+k-1} \int_y^1 (v-y)^{2j-2s-k} v^{-2j} d\varphi_{x,\theta,j}(v),$$

and $\varphi_{x,\theta,j}(y) := \mu_x q_x(\theta)\beta(d_x^+, y, j)$.

Moreover, the covariance between $\mathcal{Z}^*_{x_1,\theta_1,s}(t)$ *and* $\mathcal{Y}^*_{x_2,\theta_2}$ *is given by*

$$\mathrm{Cov}\big(\mathcal{Z}^*_{x_1,\theta_1,s}(t), \mathcal{Y}^*_{x_2,\theta_2}\big) = \beta(d_{x_1}^+, e^{-t}, s)\psi_{x_1,\theta_1,\theta_2}\mathbb{1}\{x_1 = x_2\}.$$

Proof. See Appendix 2.8.4. □

We now turn back to the proof of Theorem 2.6. For the number of solvent institutions with type $x \in \mathcal{X}$, threshold $\theta = 1, \dots, d_x^+ - 1$ and $\ell = 1, \dots, \theta - 1$ defaulted neighbors at time t, we have

$$S^{(n)}_{x,\theta,\ell}(t) = V^{(n)}_{x,\theta,d_x^+-\ell} - V^{(n)}_{x,\theta,d_x^+-\ell+1}.$$

Moreover, for $\ell = 0$, $S^{(n)}_{x,\theta,0}(t) = V^{(n)}_{x,\theta,d_x^+}$. Using the joint asymptotic normality of $V^{*(n)}_{x,\theta,d_x^+-\ell}$ and $V^{*(n)}_{x,\theta,d_x^+-\ell+1}$, we obtain that in $\mathcal{D}[0,\infty)$,

$$n^{-1/2}\left(S^{(n)}_{x,\theta,\ell}(t \wedge \tau_n) - n\mu_x^{(n)} q_x^{(n)}(\theta) b\big(d_x^+, 1 - e^{-(t\wedge\tau_n)}, \ell\big)\right) \xrightarrow{d} \mathcal{Z}_{x,\theta,\ell}(t \wedge t_0),$$

where $\mathcal{Z}_{x,\theta,\ell} = \mathcal{Z}^*_{x,\theta,d_x^+-\ell} - \mathcal{Z}^*_{x,\theta,d_x^+-\ell+1}$ for $\ell \geqslant 1$ and $\mathcal{Z}_{x,\theta,0} = \mathcal{Z}^*_{x,\theta,d_x^+}$. Further, for convenience, we set $\mathcal{Z}^*_{x,\theta,s} = 0$ for all $s > d_x^+$. Thus for any two triple (x_1, θ_1, ℓ_1) and (x_2, θ_2, ℓ_2), we have

$$\begin{aligned}
&\mathrm{Cov}\big(\mathcal{Z}_{x_1,\theta_1,\ell_1}(t), \mathcal{Z}_{x_2,\theta_2,\ell_2}(t)\big)\\
=&\mathrm{Cov}\big(\mathcal{Z}^*_{x_1,\theta_1,d_{x_1}^+-\ell_1}(t), \mathcal{Z}^*_{x_2,\theta_2,d_{x_2}^+-\ell_2}(t)\big) + \mathrm{Cov}\big(\mathcal{Z}^*_{x_1,\theta_1,d_{x_1}^+-\ell_1+1}(t), \mathcal{Z}^*_{x_2,\theta_2,d_{x_2}^+-\ell_2+1}(t)\big)\\
&- \mathrm{Cov}\big(\mathcal{Z}^*_{x_1,\theta_1,d_{x_1}^+-\ell_1}(t), \mathcal{Z}^*_{x_2,\theta_2,d_{x_2}^+-\ell_2+1}(t)\big) - \mathrm{Cov}\big(\mathcal{Z}^*_{x_1,\theta_1,d_{x_1}^+-\ell_1+1}(t), \mathcal{Z}^*_{x_2,\theta_2,d_{x_2}^+-\ell_2}(t)\big),
\end{aligned} \tag{2.11}$$

where the covariances on the right hand side can be computed by using Lemma 2.14.

In particular, the variance of $\mathcal{Z}_{x,\theta,\ell}(t)$ is given by

$$\begin{aligned}
\sigma_{x,\theta,\ell}(t) =&\tilde{\sigma}_{x,\theta,d_x^+-\ell,d_x^+-\ell}(e^{-t}) + \tilde{\sigma}_{x,\theta,d_x^+-\ell+1,d_x^+-\ell+1}(e^{-t})\\
&- 2\tilde{\sigma}_{x,\theta,d_x^+-\ell,d_x^+-\ell+1}(e^{-t}) + b^2(d_x^+, e^{-t}, d_x^+ - \ell)\psi_{x,\theta,\theta},
\end{aligned} \tag{2.12}$$

where $\psi_{x,\theta,\theta} = \mu_x q_x(\theta)(1 - q_x(\theta))$ as in Lemma 2.14.

2.7.4 Proof of Theorem 2.7

We first state a lemma which holds under the moment regularity condition (i.e., Assumption 2.3b).

Lemma 2.15. *The Assumption 2.3b guarantees that, as* $n \to \infty$, $f^{(n)}_{\clubsuit}(z) \to f_{\clubsuit}(z)$, *for all* $\clubsuit \in \{S, D, H^+, I^+, W\}$, *together with all their derivatives, uniformly on* $[0, 1]$.

Proof. See Appendix 2.8.5. □

This lemma allows us to extend the convergence results to some infinite sums of $V^{*(n)}_{x,\theta,s}(t)$. We denote by $\mathcal{X}^+_\ell$ and $\mathcal{X}^-_\ell$ the set of characteristics $x \in \mathcal{X}$ with in-degree $d^+_x \geqslant \ell$ and out-degree $d^-_x \geqslant \ell$, respectively.

Lemma 2.16. *Let $\tau_n \leqslant \tau^\star_n$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Under Assumption 2.3b and for n large enough, we have as $\ell \to \infty$,*

$$\mathbb{E}\Big[\sup_{t \leqslant T}\Big|\sum_{x \in \mathcal{X}^+_\ell}\sum_{\theta=1}^{d^+_x}\sum_{s=d^+_x-\theta+2}^{d^+_x} V^{*(n)}_{x,\theta,s}(t)\Big|\Big] \to 0,$$

Further as $n \to \infty$,

$$\sum_{x \in \mathcal{X}}\sum_{\theta=1}^{d^+_x}\sum_{s=d^+_x-\theta+2}^{d^+_x} V^{*(n)}_{x,\theta,s}(t \wedge \tau_n) \xrightarrow{d} \sum_{x \in \mathcal{X}}\sum_{\theta=1}^{d^+_x}\sum_{s=d^+_x-\theta+2}^{d^+_x} \mathcal{Z}^*_{x,\theta,s}(t \wedge t_0). \tag{2.13}$$

Similarly, as $n \to \infty$,

$$\sum_{x \in \mathcal{X}}\sum_{\theta=1}^{d^+_x} V^{*(n)}_{x,\theta,d^+_x-\theta+1}(t \wedge \tau_n) \xrightarrow{d} \sum_{x \in \mathcal{X}}\sum_{\theta=1}^{d^+_x} \mathcal{Z}^*_{x,\theta,d^+_x-\theta+1}(t \wedge t_0),$$

$$\sum_{x \in \mathcal{X}}\sum_{\theta=1}^{d^+_x}(d^+_x - \theta + 1) V^{*(n)}_{x,\theta,d^+_x-\theta+1}(t \wedge \tau_n) \xrightarrow{d} \sum_{x \in \mathcal{X}}\sum_{\theta=1}^{d^+_x}(d^+_x - \theta + 1) \mathcal{Z}^*_{x,\theta,d^+_x-\theta+1}(t \wedge t_0),$$

and

$$\sum_{x \in \mathcal{X}} d^-_x \sum_{\theta=1}^{d^+_x} V^{*(n)}_{x,\theta,d^+_x-\theta+1}(t \wedge \tau_n) \xrightarrow{d} \sum_{x \in \mathcal{X}} d^-_x \sum_{\theta=1}^{d^+_x} \mathcal{Z}^*_{x,\theta,d^+_x-\theta+1}(t \wedge t_0).$$

Moreover, all the above limit processes on the right hand side are continuous.

Proof. see Appendix 2.8.6. □

We come back to the proof of Theorem 2.7. Recall that $L_n(t)$ denotes the total number of alive in balls at time t. At the initial time, $L_n(0) = n\lambda^{(n)}$ and $L_n(t)$ decreases by 1 each time a (in) ball dies. Since the death happens after an exponential time with rate 1 independently, therefore on $[0, \tau^\star_n]$, writing in differential form, we have

$$dL_n(t) = -L_n(t)dt + d\mathcal{M}_t,$$

where $\mathcal{M}$ is a martingale.

Then by similar arguments as in the proof of [152, Theorem 3.1] for W_n, with obvious change of the jump from -2 to -1, we obtain

$$n^{-1/2}\big(L_n(t \wedge \tau_n) - L_n(0)\big) \xrightarrow{d} \mathcal{Z}_L(t \wedge t_0) \qquad \text{in} \quad \mathcal{D}\,[0, \infty)\,,$$

where $\mathcal{Z}_L$ is a continuous Gaussian process with $\mathbb{E}[\mathcal{Z}_L(t)] = 0$ and covariances

$$\mathbb{E}\big[\mathcal{Z}_L(t)\mathcal{Z}_L(u)\big] = \lambda(e^{-t} - e^{-2t})/2, \qquad 0 \leqslant t \leqslant u < \infty.$$

Let $U^{(n)}_{x,\theta,s}(t)$ and $V^{(n)}_{x,\theta,s}(t)$ be as defined in the proof of Theorem 2.7.3. Note that

$$S_n(t) = \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} V^{(n)}_{x,\theta,d_x^+-\theta+1}(t), \quad D_n(t) = n - S_n(t), \tag{2.14}$$

and,

$$H_n^+(t) = \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \sum_{s=d_x^+-\theta+1}^{d_x^+} sU^{(n)}_{x,\theta,s}(t) = \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \Big[(d_x^+ - \theta + 1)V^{(n)}_{x,\theta,d_x^+-\theta+1}(t) + \sum_{s=d_x^+-\theta+2}^{d_x^+} V^{(n)}_{x,\theta,s}(t)\Big]. \tag{2.15}$$

Further, $I_n^+(t) = L_n(t) - H_n^+(t)$, and,

$$W_n(t) = L_n(t) - \sum_{x\in\mathcal{X}} d_x^- \sum_{\theta=1}^{d_x^+} V^{(n)}_{x,\theta,d_x^+-\theta+1}(t). \tag{2.16}$$

Then by Lemma 2.14, combining (2.14), (2.15), (2.16) and the convergences results for the infinite sums in Lemma 2.16, we obtain that for $\clubsuit \in \{S, H^+, I^+, W\}$,

$$n^{-1/2}\left(\clubsuit_n(t\wedge\tau_n) - n\widehat{f}^{(n)}_{\clubsuit}(t\wedge\tau_n)\right) \xrightarrow{d} \mathcal{Z}_{\clubsuit}(t\wedge t_0),$$

with

$$\mathcal{Z}_S := \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \mathcal{Z}^*_{x,\theta,d_x^+-\theta+1}, \tag{2.17}$$

$$\mathcal{Z}_{H^+} := \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} (d_x^+ - \theta + 1)\mathcal{Z}^*_{x,\theta,d_x^+-\theta+1} + \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \sum_{s=d_x^+-\theta+2}^{d_x^+} \mathcal{Z}^*_{x,\theta,s}, \tag{2.18}$$

$$\mathcal{Z}_{I^+} := \mathcal{Z}_L - \mathcal{Z}_{H^+}, \tag{2.19}$$

and

$$\mathcal{Z}_W := \mathcal{Z}_L - \sum_{x\in\mathcal{X}} d_x^- \sum_{\theta=1}^{d_x^+} \mathcal{Z}^*_{x,\theta,d_x^+-\theta+1}. \tag{2.20}$$

Hence we have proved the asymptotic normality in Theorem 2.7. The covariances between all the processes are given by (2.55)-(2.60) in Appendix 2.8.7.

2.7.5 Proof of Theorem 2.8

Since $\alpha > 0$, for a positive constant ε small enough, we have $f_W(z^\star - \varepsilon) < 0$ and $f_W(z^\star + \varepsilon) > 0$. By Lemma 2.15, we have $f_W^{(n)} \to f_W$ uniformly on $[0,1]$. For n large enough, we also have $f_W^{(n)}(z^\star - \varepsilon) < 0$ and $f_W^{(n)}(z^\star + \varepsilon) > 0$. Hence for n large enough, there exists a sequence $\widehat{z}_n$ in $(z^\star - \varepsilon, z^\star + \varepsilon)$ such that $f_W^{(n)}(\widehat{z}_n) = 0$ and $f_W^{(n)} > 0$ on $[z^\star + \varepsilon, 1]$. Since ε can be arbitrarily small, we obtain $\widehat{z}_n \to z^\star$. Define $\widehat{t}_n := -\ln \widehat{z}_n$. Consequently we also have $\widehat{t}_n \to t^\star$.

By using the Skorokhod's representation theorem [156, Theorem 3.30], we can change the probability space so that all the random variables are well defined and all the convergence results of Theorem 2.7 and $\tau_n^\star \to t^\star$ (from Lemma 2.2) hold a.s..

Taking $t = \tau_n^\star$ and $t_0 = t^\star$, we get

$$W_n(\tau_n^\star) = n\widehat{f}_W^{(n)}(\tau_n^\star) + n^{1/2}\mathcal{Z}_W(\tau_n^\star \wedge t^\star) + o(n^{1/2}) = n\widehat{f}_W^{(n)}(\tau_n^\star) + n^{1/2}\mathcal{Z}_W(t^\star) + o(n^{1/2}),$$

by the continuity of $\mathcal{Z}_W$. Since $W_n(\tau_n^\star) = -1$, then $\widehat{f}_W^{(n)}(\tau_n^\star) = -n^{-1/2}\mathcal{Z}_W\big(t^\star\big) + o(n^{-1/2})$. Since, as $n \to \infty$, $\tau_n^\star \to t^\star$ and $\widehat{t}_n \to t^\star$ hold a.s., there exists some ξ_n in the interval between $\widehat{t}_n$ and $\tau_n^\star$ such that $\xi_n \to t^\star$. Further, by Lemma 2.15,

$$(\widehat{f}_W^{(n)})'(\xi_n) \to \widehat{f}_W'(t^\star) = -z^\star\alpha.$$

It follows then by Mean-Value theorem that

$$\widehat{f}_W^{(n)}(\tau_n^\star) = \widehat{f}_W^{(n)}(\tau_n^\star) - \widehat{f}_W^{(n)}(\widehat{t}_n) = (\widehat{f}_W^{(n)})'(\xi_n)(\tau_n^\star - \widehat{t}_n) = (-z^\star\alpha + o(1))(\tau_n^\star - \widehat{t}_n).$$

Hence we have

$$\tau_n^\star - \widehat{t}_n = \Big(-\frac{1}{z^\star\alpha} + o(1)\Big)\widehat{f}_W^{(n)}(\tau_n^\star) = n^{-1/2}\frac{1}{z^\star\alpha}(\mathcal{Z}_W(t^\star) + o(1)).$$

Then, by a similar argument for $S_n(\tau_n^\star)$, combining the above formula and Lemma 2.15, we have a.s. for some $\xi_n' \to t^\star$,

$$\begin{aligned} n^{-1/2}S_n(\tau_n^\star) &= n^{1/2}\widehat{f}_S^{(n)}(\tau_n^\star) + \mathcal{Z}_S(t^\star) + o(1) \\ &= n^{1/2}\widehat{f}_S^{(n)}(\tau_n^\star) + n^{1/2}(\widehat{f}_S^{(n)})'(\xi_n')(\tau_n^\star - \widehat{t}_n) + \mathcal{Z}_S(t^\star) + o(1) \\ &= n^{1/2}f_S^{(n)}(\widehat{z}_n) + \frac{\widehat{f}_S'(t^\star)}{\alpha z^\star}\mathcal{Z}_W(t^\star) + \mathcal{Z}_S(t^\star) + o(1) \\ &= n^{1/2}f_S^{(n)}(\widehat{z}_n) - \frac{f_S'(z^\star)}{\alpha}\mathcal{Z}_W(t^\star) + \mathcal{Z}_S(t^\star) + o(1), \end{aligned} \tag{2.21}$$

where the last equality follows from the fact that $(\widehat{f}_S)'(t) = -(f_S)(e^{-t})e^{-t}$ and $e^{-t^\star} = z^\star$.

Using similar arguments, we have the other analogues in Theorem 2.8.

2.7.6 Proof of Theorem 2.9

Using the same notation as in the proof of Theorem 2.1 (see Section 2.7.1), we have $S^{(n)}_{x,\theta,\ell}(t) = U^{(n)}_{x,\theta,d_x^+-\ell}(t)$. Then we have

$$\sum_{x\in\mathcal{X}} \bar{L}^{\diamondsuit}_x \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell S^{(n)}_{x,\theta,\ell}(t) = \sum_{x\in\mathcal{X}} \bar{L}^{\diamondsuit}_x \sum_{\theta=1}^{d_x^+} \sum_{\ell=1}^{\theta-1} \ell U^{(n)}_{x,\theta,d_x^+-\ell}(t),$$

and,

$$\sum_{x\in\mathcal{X}} \bar{L}^{\odot}_x D^{(n)}_x(t) = \sum_{x\in\mathcal{X}} \bar{L}^{\odot}_x (n\mu^{(n)}_x - \sum_{\theta=1}^{d_x^+} \sum_{s=d_x^+-\theta+1}^{d_x^+} U^{(n)}_{x,\theta,s}(t)).$$

Since for all $x \in \mathcal{X}$, $\bar{L}^{\odot}_x$ and $\bar{L}^{\diamondsuit}_x$ are bounded, there exists some constant C such that the two above expressions are both dominated by

$$nC \sum_{x\in\mathcal{X}} (d_x^+ + d_x^-) \sum_{\theta=1}^{d_x^+} \sum_{s=d_x^+-\theta+1}^{d_x^+} U^{(n)}_{x,\theta,s}(t).$$

Thus again by Lemma 3.22 and basic computations, it follows that

$$\sup_{t\leqslant\tau_n} \Big|\frac{\Gamma^{\diamondsuit}_n(t)}{n} - f_{\diamondsuit}(e^{-t})\Big| \xrightarrow{p} 0.$$

For $z^\star = 0$, by Lemma 2.2, $\tau_n^\star \xrightarrow{p} \infty$. Then we have $e^{-\tau_n^\star} \xrightarrow{p} 0$ and $f_{\diamondsuit}(0) = \bar{\Gamma}^{\diamondsuit} - \sum_{x\in\mathcal{X}} \bar{L}^{\odot}_x \mu_x$. By the continuity of $f_{\diamondsuit}$, it follows that

$$f_{\diamondsuit}(e^{-\tau_n^\star}) = \bar{\Gamma}^{\diamondsuit} - \sum_{x\in\mathcal{X}} \bar{L}^{\odot}_x \mu_x + o_p(1).$$

We therefore have

$$\frac{\Gamma^{\diamondsuit}_n(\tau_n^\star)}{n} \xrightarrow{p} \bar{\Gamma}^{\diamondsuit} - \sum_{x\in\mathcal{X}} \bar{L}^{\odot}_x \mu_x.$$

For $z^\star \in (0,1]$ and $f'_W(z^\star) > 0$, by Lemma 2.2, $\tau_n^\star \xrightarrow{p} -\ln z^\star$. Then a similar argument as above implies that

$$\frac{\Gamma^{\diamondsuit}_n(\tau_n^\star)}{n} \xrightarrow{p} f_{\diamondsuit}(z^\star).$$

Recall that we have defined $V^{(n)}_{x,\theta,s}(t)$ as the number of bins (institutions) with type $x \in \mathcal{X}$, threshold θ and at least s alive balls at time t. We notice that

$$\sum_{x\in\mathcal{X}} \bar{L}^{\odot}_x D^{(n)}_x(t) = \sum_{x\in\mathcal{X}} \bar{L}^{\odot}_x \big(n\mu^{(n)}_x - \sum_{\theta=1}^{d_x^+} V^{(n)}_{x,\theta,d_x^+-\theta+1}(t)\big),$$

and,

$$\sum_{\ell=1}^{\theta-1} \ell S_{x,\theta,\ell}^{(n)}(t) = (\theta-1)V_{x,\theta,d_x^+-\theta+1}^{(n)}(t) - \sum_{s=d_x^+-\theta+2}^{d_x^+} V_{x,\theta,s}^{(n)}(t).$$

Since for all $x \in \mathcal{X}$, $\bar{L}_x^{\odot}$ and $\bar{L}_x^{\lozenge}$ are bounded, there exists a constant C such that $\bar{L}_x^{\odot} + \bar{L}_x^{\lozenge} \leqslant C$ for all $x \in \mathcal{X}$. Hence, by using similar arguments as in Appendix 2.7.4 to prove the convergence of H_n^+, S_n and so on, Lemma 2.16 leads to the convergence of the following (infinite) sums

$$\sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} \sum_{\theta=1}^{d_x^+} V_{x,\theta,d_x^+-\theta+1}^{*(n)}(t) \xrightarrow{d} \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} \sum_{\theta=1}^{d_x^+} \mathcal{Z}_{x,\theta,d_x^+-\theta+1}^{*}(t),$$

and

$$\sum_{x\in\mathcal{X}} \bar{L}_x^{\lozenge} \sum_{\theta=1}^{d_x^+} [(\theta-1)V_{x,\theta,d_x^+-\theta+1}^{*(n)}(t) - \sum_{s=d_x^+-\theta+2}^{d_x^+} V_{x,\theta,s}^{*(n)}(t)]$$
$$\xrightarrow{d} \sum_{x\in\mathcal{X}} \bar{L}_x^{\lozenge} \sum_{\theta=1}^{d_x^+} [(\theta-1)\mathcal{Z}_{x,\theta,d_x^+-\theta+1}^{*}(t) - \sum_{s=d_x^+-\theta+2}^{d_x^+} \mathcal{Z}_{x,\theta,s}^{*}(t)].$$

Hence we have in $\mathcal{D}[0,\infty)$, as $n \to \infty$,

$$n^{-1/2}\left(\Gamma_n^{\lozenge}(t\wedge\tau_n) - n\widehat{f}_{\lozenge}^{(n)}(t\wedge\tau_n)\right) \xrightarrow{d} \mathcal{Z}_{\lozenge}(t\wedge t_0), \tag{2.22}$$

where $\mathcal{Z}_{\lozenge}$ is a continuous Gaussian process with,

$$\mathcal{Z}_{\lozenge} := -\sum_{x\in\mathcal{X}} \bar{L}_x^{\lozenge} \sum_{\theta=1}^{d_x^+}\Big[(\theta-1)\mathcal{Z}_{x,\theta,d_x^+-\theta+1}^{*}(t) - \sum_{s=d_x^+-\theta+2}^{d_x^+} \mathcal{Z}_{x,\theta,s}^{*}(t)\Big] - \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} \sum_{\theta=1}^{d_x^+} \mathcal{Z}_{x,\theta,d_x^+-\theta+1}^{*}(t). \tag{2.23}$$

Note that the convergence result (2.22) holds jointly with the convergence of other processes H_n^+, S_n and so on. Hence, by a similar argument as in Section 2.7.5, we have that

$$n^{-1/2}\Gamma_n^{\lozenge}(\tau_n^{\star}) = n^{1/2} f_{\lozenge}^{(n)}(\widehat{z}_n) - \frac{f_{\lozenge}'(z^{\star})}{\alpha}\mathcal{Z}_W(t^{\star}) + \mathcal{Z}_{\lozenge}(t^{\star}) + o(1).$$

This gives

$$\mathcal{Z}_{\lozenge}^{\star} := \mathcal{Z}_{\lozenge}(t^{\star}) - \alpha^{-1} f_{\lozenge}'(z^{\star})\mathcal{Z}_W(t^{\star}) = \mathcal{Z}_{\lozenge}(t^{\star}) - \Delta(z^{\star})\mathcal{Z}_W(t^{\star}).$$

Hence $\mathcal{Z}_{\lozenge}^{\star}$ is a centered Gaussian random variable, which completes the proof.

2.7.7 Proof of Theorem 2.10

The proof is based on Theorem 2.1 and some arguments of [150] used to study the conditions for existence of giant component in the percolated random (non directed) graph with given vertex degrees.

We first remove all potential saved links by the planner from the network. Consider the type-dependent bond percolation model where we remove each incoming link to any institution of type $x \in \mathcal{X}$ with probability α_x. Note that this also includes extra removed links between solvent institutions that will not play any role in the default contagion process. Next we run the death process as described in Section 2.2.3 and Appendix 2.7.1.

Let $\widetilde{W}_n(t)$ and $\widetilde{D}_n(t)$ denote respectively the number of white balls and the total number of defaults at time t in the percolated random graph. The surviving probability for each ball of type x at time t is $\alpha_x + (1-\alpha_x)e^{-t}$ and by following the same steps as for the proof of Theorem 2.1 in Section 2.7.1, we obtain

$$\sup_{t\leqslant\tau_n}\Big|\frac{\widetilde{W}_n(t)}{n} - f_W^{(\boldsymbol{\alpha})}(e^{-t})\Big| \xrightarrow{p} 0, \quad \sup_{t\leqslant\tau_n}\Big|\frac{\widetilde{D}_n(t)}{n} - f_D^{(\boldsymbol{\alpha})}(e^{-t})\Big| \xrightarrow{p} 0.$$

Let $\tilde{\tau}_n^\star$ be the first time when $\widetilde{W}_n(\tilde{\tau}_n^\star) = -1$. Then, similar to the proof of Lemma 2.2 in Section 2.8.2, we find that $\tilde{\tau}_n^\star \to -\ln z_{\boldsymbol{\alpha}}^\star$, where $z_{\boldsymbol{\alpha}}^\star := \sup\{z \in [0,1] : f_W^{(\boldsymbol{\alpha})}(z) = 0\}$ and

$$f_W^{(\boldsymbol{\alpha})}(z) := \lambda z - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=1}^{d_x^+} q_x(\theta)\beta\big(d_x^+, \alpha_x + (1-\alpha_x)z, d_x^+ - \theta + 1\big).$$

Next, by following the same steps as proof of Theorem 2.4 in Section 2.7.2, we obtain

$$\frac{S_{x,\theta,\ell}^{(n)}(\boldsymbol{\alpha}_n)}{n} \xrightarrow{p} s_{x,\theta,\ell}^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star), \qquad \frac{D_n(\boldsymbol{\alpha}_n)}{n} \xrightarrow{p} f_D^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star),$$

which then implies (by definition) that the system-wide wealth converges to

$$\frac{\Gamma_n^\Diamond(\boldsymbol{\alpha}_n)}{n} \xrightarrow{p} f_\Diamond^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star) := \bar{\Gamma}^\Diamond - \sum_{x\in\mathcal{X}} \bar{L}_x^\odot f_D^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star) - \sum_{x\in\mathcal{X}} \bar{L}_x^\Diamond \sum_{\theta=1}^{d_x^+}\sum_{\ell=1}^{\theta-1} \ell s_{x,\theta,\ell}^{(\boldsymbol{\alpha})}(z_{\boldsymbol{\alpha}}^\star).$$

The total cost of interventions $\boldsymbol{\alpha}_n$ for the planner converges to

$$\frac{\Phi_n(\boldsymbol{\alpha}_n)}{n} \xrightarrow{p} \phi(z_{\boldsymbol{\alpha}}^\star) := \sum_{x\in\mathcal{X}} \mu_x\alpha_x C_x \sum_{\ell=1}^{d_x^+} \ell b\left(d_x^+, 1 - z_{\boldsymbol{\alpha}}^\star, \ell\right).$$

This completes the proof of Theorem 2.10.

2.7.8 Proofs in §2.6

Proof of Theorem 2.11

The main idea of proof is to alter the speed of default cascade process by multiplying each transition rate by a constant depending on the current state and transfer the default cascade process to its associated death process in §2.2.3.

Recall that all agents (at any time t) are of two states and all half-edges are of four different types; these are denoted by $\mathbf{S}(t)$ (solvent), $\mathbf{D}(t)$ (defaulted) agents, and further $\mathbf{H}^+(t)$ (healthy in), $\mathbf{H}^-(t)$ (healthy out), $\mathbf{I}^+(t)$ (infected in) and $\mathbf{I}^-(t)$ (infected out) half-edges, respectively. We also let $\mathbf{B}^+(t) := \mathbf{H}^+(t) \cup \mathbf{I}^+(t)$ be the total in half-edges (balls) at time t. The above quantities without bold format dontes the total number of corresponding objects at time t, e.g., $B^+(t) = |\mathbf{B}^+(t)|$ denotes the total number of remaining in half-edges at time t. In particular, $W(t)$ denotes the number of $\mathbf{I}^-$ half-edges (white balls) at time t and controls the stopping time of our default cascade process: $\tau_n^\star := \inf\{t \geqslant 0 : W_n(t) = 0\}$.

Now note that to run the default cascade process in its associated death process in §2.2.3, we need to change the transition rate. Namely, if there are x_W white balls ($\mathbf{I}^-$ half-edges) and x_B in balls (half-edges) in total, we multiply all transition rates out of such a state by x_B/x_I. This means that if at time t, there are $W(t)$ white balls remaining, we do not wait an exponential time with mean $W(t)$ to reveal next infected link, but instead we wait an exponential time with mean $B^+(t)$ for the next reveal. Thus the new reveal rate will be determined by $B^+(t)$ and the reveal point process is a Poisson process with rate $B^+(t)$. Hence it will be equivalent to the death process in §2.2.3, where every in half-edge of any type pair with a random infected out half-edge uniformly random and independently after an exponential time with mean one, until all there are no more infected out half-edges.

We next recover the asymptotic behavior in the default cascade process from its associated death process (Theorem 2.1) by using the following result for time changed Markov chains from [153, Lemma A.1]; see also [184, III. (21.7)] for a more general result.

Lemma 2.17. *Suppose $(Y_t)_{t\geqslant 0}$ is a continuous time Markov chain with finite state space E and infinitesimal transition rates $(q(i,j))_{i,j\in E}$. Let $f : E \to (0,\infty)$ be strictly positive and define the strictly increasing process*

$$A_\tau = \int_0^\tau f(Y_s)ds \qquad \tau \geqslant 0,$$

and its inverse $\tau(t)$, $t \geqslant 0$. Then the time-changed process $(Y_{\tau(t)})_{t\geqslant 0}$ is again Markov and has infinitesimal rates $(q(i,j)/f(i))_{i,j\in E}$.

Hence, in order to study the default cascade process from its associated death process, we need to use Lemma 2.17 by setting

$$A_\tau^{(n)} = \int_0^\tau \frac{\widetilde{B}_n^+(s)}{\widetilde{W}_n(s)}ds, \qquad \tau \geqslant 0,$$

where for $s \geqslant \widetilde{\tau}_n^\star$, we set the term in the integrand (i.e., $\widetilde{B}_n^+(s)/\widetilde{W}_n(s)$) to 1.

Then $A^{(n)}$ is a continuous strictly increasing function. We denote by $\tau_n(t) : [0,\infty) \to [0,\infty)$ the strictly increasing continuous inverse of $A^{(n)}$ such that $A_{\tau_n(t)}^{(n)} = t$ for any $t \geqslant 0$.

Hence, by Lemma 2.17, the stochastic processes in the original default cascade process could be realized by setting (for all $t \geqslant 0$) $D_n(t) = \widetilde{D}_n(\tau_n(t)), I_n^+(t) = \widetilde{I}_n^+(\tau_n(t))$, and the same for all other processes. Since the default cascade process stops when we reveal all infected out half-edges, we replace $\tau_n(t)$ by $\widehat{\tau}_n(t) := \tau_n(t) \wedge \widetilde{\tau}_n^\star$.

On the other hand, since in the time-changed process, every in half edge has an i.i.d. exponential lifetime with parameter one, we have, by using the Glivenko-Cantelli theorem, as $n \to \infty$,

$$\sup_{t \geqslant 0} \left| \frac{\widetilde{B}_n^+(t)}{n\lambda^{(n)}} - e^{-t} \right| \xrightarrow{p} 0. \tag{2.24}$$

Moreover, by Theorem 2.1, the limit function of $\widetilde{W}_n(t)$ is $f_W(e^{-t})$. In this sense, we define the following strictly increasing function A_τ by

$$A_\tau := \int_0^\tau \frac{\lambda e^{-s}}{f_W(e^{-s})} ds, \qquad 0 \leqslant \tau < -\ln z^\star, \tag{2.25}$$

which can be regarded as the limit function of $A_\tau^{(n)}$ as $n \to \infty$.

We are now ready to prove point (i) of Theorem 2.11. We denote by $\widehat{\tau}(t)$ the inverse of A and set $v(t) := e^{-\widehat{\tau}(t)}$. We show that $v(t)$ satisfies point (i).

Note that the integrand in (2.25) is strictly positive. Further, for sufficiently small $\epsilon > 0$, and for all $x \in [z^\star, z^\star + \epsilon]$, we have $f_W(x) < \lambda(x - z^\star)$. So when $e^{-\tau}$ is sufficiently close to $z^\star$, $\lambda e^{-\tau}/f_W(e^{-\tau}) > e^{-\tau}/(e^{-\tau} - z^\star)$, and it follows that $A_\tau \nearrow \infty$ as $\tau \to -\ln z^\star$ (for $z^\star = 0$, as $\tau \to \infty$). Moreover, the inverse $\widehat{\tau}$ is clearly strictly increasing and continuous differentiable. By the Inverse Function Theorem, we have $\widehat{\tau}'(t) = f_W(e^{-\widehat{\tau}(t)})/\lambda e^{-\widehat{\tau}(t)}$. Therefore $v(t)$ satisfies (ODE) with initial value $v(0) = 1$.

On the other hand, the coefficient of (ODE) is Lipschitz continuous on $[0, 1]$ by Assumption 2.4. Indeed, for $1 \leqslant \ell \leqslant d - 1$, we have

$$\frac{\partial b(d, z, \ell)}{\partial z} = db(d-1, z, \ell - 1) - db(d-1, z, r),$$

and

$$\frac{\partial b(d, z, 0)}{\partial z} = -db(d-1, z, 0), \qquad \frac{\partial b(d, z, d)}{\partial z} = db(d-1, z, d-1).$$

We thus have for $0 < \ell \leqslant d$,

$$\frac{\partial \beta(d, z, \ell)}{\partial z} = db(d-1, z, \ell - 1).$$

It follows then by Assumption 2.4,

$$\begin{aligned} f_W'(z) &\leqslant \lambda + \sum_{x \in \mathcal{X}} \mu_x d_x^- \sum_{\theta=1}^{d_x^+} q_x(\theta) d_x^+ b(d_x^+, z, d_x^+ - \theta) \\ &\leqslant \lambda + \sum_{x \in \mathcal{X}} \mu_x d_x^- d_x^+ \leqslant \lambda + \sum_{x \in \mathcal{X}} \mu_x (d_x^- + d_x^+)^2 \leqslant C, \end{aligned}$$

for some (sufficiently large) constant C. The existence and uniqueness of the solution $v(t)$ is now guaranteed. Note that the constant function $z^\star$ is also a solution to (ODE). By the Comparison Theorem, there exists a solution $v(t)$ such that $v(t) > z^\star$ for all $t \geqslant 0$. Since $v(t)$ is strictly decreasing and bounded from below by $z^\star$, it must converge to $z^\star$. The proof for point (i) is now complete.

We now show that $v(t)$ is the limit of $e^{-\widehat{\tau}_n(t)}$, for $t \geqslant 0$, as $n \to \infty$. First note that for some small $\epsilon > 0$, for all $s \in [0, -\ln z^\star - \epsilon]$, we have from Theorem 2.1, (uniformly)

$$\frac{\widetilde{B}_n^+(s)}{\widetilde{W}_n(s)} \xrightarrow{p} \frac{\lambda e^{-s}}{f_W(e^{-s})}.$$

By the definition of $z^\star$, if $z^\star > 0$, there exists some small δ such that $f_W(e^{-s}) \geqslant \delta > 0$ for all $s \in [0, -\ln z^\star - \epsilon]$. It follows thus, by the Dominated Convergence Theorem,

$$A_\tau^{(n)} \xrightarrow{p} A_\tau, \tag{2.26}$$

uniformly on $[0, -\ln z^\star - \epsilon]$. If $z^\star = 0$, for any fixed t_0, the above convergence holds uniformly on $[0, t_0]$.

We first analyze the case where $z^\star > 0$. Let $t_1 := A_{-\ln z^\star - 2\epsilon}$. Since A is strictly increasing, we have for n large enough, w.h.p. $A^{(n)}_{-\ln z^\star - \epsilon} > t_1$. So for $t \leqslant t_1$, it follows that w.h.p. $\widehat{\tau}_n(t) < -\ln z^\star - \epsilon$. Hence by (2.26) we have

$$A_{\widehat{\tau}_n(t)} \xrightarrow{p} t = A_{\widehat{\tau}(t)}, \tag{2.27}$$

uniformly on $[0, t_1]$. Recall that $\widehat{\tau}'(t) = f_W(e^{-\widehat{\tau}})/\lambda e^{-\widehat{\tau}} \leqslant 1$, which implies that $\widehat{\tau}$ is uniformly continuous. Combining with (2.27), we have

$$\sup_{t \leqslant t_1} |\widehat{\tau}_n(t) - \widehat{\tau}(t)| \xrightarrow{p} 0. \tag{2.28}$$

By Lemma 2.2, we have $-\ln z^\star - \epsilon < \widetilde{\tau}_n^\star < -\ln z^\star + \epsilon$. Then it follows that w.h.p. for $t \geqslant t_1$ and n large enough,

$$-\ln z^\star - 3\epsilon \leqslant \widehat{\tau}(t_1) - \epsilon = \widehat{\tau}_n(t_1) \leqslant \widehat{\tau}_n(t) \leqslant \widetilde{\tau}_n^\star < -\ln z^\star + \epsilon.$$

Further, for $t \geqslant t_1$,

$$-\ln z^\star - 2\epsilon \leqslant \widehat{\tau}(t_1) \leqslant \widehat{\tau}(t) < -\ln z^\star + \epsilon.$$

So, by taking $\epsilon \to 0$, we have

$$\sup_{t \geqslant t_1} |\widehat{\tau}_n(t) - \widehat{\tau}(t)| \xrightarrow{p} 0. \tag{2.29}$$

We thus obtain $\widehat{\tau}_n(t) \xrightarrow{p} \widehat{\tau}(t)$ uniformly for all $t \geqslant 0$, and so $e^{-\widehat{\tau}_n(t)} \xrightarrow{p} e^{-\widehat{\tau}(t)} = v(t)$ uniformly for all $t \geqslant 0$. It therefore follows that

$$\begin{aligned}\sup_{t\geqslant 0} \Big|\frac{S_n(t)}{n} - f_S(v(t))\Big| &= \sup_{t \geqslant 0} \Big|\frac{\widetilde{S}_n(\widehat{\tau}_n(t))}{n} - f_S(v(t))\Big| \\ &\leqslant \sup_{t\geqslant 0} \Big|\frac{\widetilde{S}_n(\widehat{\tau}_n(t))}{n} - f_S(e^{-\widehat{\tau}_n(t)})\Big| + \sup_{t \geqslant 0} |f_S(e^{-\widehat{\tau}_n(t)}) - f_S(v(t))|,\end{aligned}$$

which converges to 0 in probability. Indeed, the first term converges to 0 by Theorem 2.1 and the second term converges to 0 by the uniform continuity of f_S on $[0, 1]$. The other convergence results for $D_n(t)$ and $W_n(t)$ follow similarly.

We now consider the case $z^\star = 0$. We have for any fixed $t_0 > 0$, $A_\tau^{(n)} \xrightarrow{p} A_\tau$ uniformly on $[0, t_0]$. By similar arguments as above, letting $t_1 = t_0 - \epsilon$ for some small $\epsilon > 0$, we have

$$\sup_{t \leqslant t_1} |\hat{\tau}_n(t) - \hat{\tau}(t)| \xrightarrow{p} 0, \tag{2.30}$$

For small $\epsilon > 0$, we can choose t_1 large enough such that both $e^{-\hat{\tau}_n(t)} < \epsilon$ and $e^{-\hat{\tau}(t)} < \epsilon$ for all $t \geqslant t_1$. Then we have

$$\begin{aligned}\sup_{t \geqslant 0} |\frac{S_n(t)}{n} - f_S(v(t))| \leqslant \sup_{t \geqslant 0} |\frac{\widetilde{S}_n(\hat{\tau}_n(t))}{n} - f_S(e^{-\hat{\tau}_n(t)})| + \sup_{t \leqslant t_1} |f_S(e^{-\hat{\tau}_n(t)}) - f_S(v(t))| \\ + \sup_{t \geqslant t_1} |f_S(e^{-\hat{\tau}_n(t)}) - f_S(v(t))|,\end{aligned}$$

which also converges to 0 in probability. The first two terms converges to 0 by the same arguments as above. For the last term, by the uniform continuity of f_S, we can take ϵ arbitrarily small to conclude. The other convergence results for $D_n(t)$ and $W_n(t)$ are similar. The proof of point (ii) is now complete and therefore, this completes the proof of Theorem 2.11.

Proof of Theorem 2.12

Note that $L_n^\star = n\lambda^{(n)} - B_n^+(\tau_n^\star)$. So, by Equation 2.24 and Assumption 2.4, as $n \to \infty$, we have $L_n^\star/n \xrightarrow{p} \lambda(1 - z^\star)$.

Let $\widetilde{T}_n(\ell)$ be the (random) time that the financial network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ faces its ℓ-th loss in the time-changed death process in §2.2.3. By Equation 2.24 and using the monotonicity and continuity of e^{-t} on t, we have for all $0 \leqslant a < 1 - z^\star$,

$$\widetilde{T}_n(a\lambda^{(n)}n) \xrightarrow{p} -\ln(1 - a). \tag{2.31}$$

Then $T_n(a\lambda^{(n)}n)$ is just the corresponding time of $\widetilde{T}_n(a\lambda^{(n)}n)$ in the original process. Thus by the time change rule, as in the proof of Theorem 2.11, we have

$$T_n(a\lambda^{(n)}n) = A^{(n)}_{\widetilde{T}_n(a\lambda^{(n)}n)} = \int_0^{\widetilde{T}_n(a\lambda^{(n)}n)} \frac{\widetilde{B}_n^+(s)}{\widetilde{W}_n(s)} ds.$$

On the other hand, we have the following decomposition:

$$\begin{aligned}&|\int_0^{\widetilde{T}_n(a\lambda^{(n)}n)} \frac{\widetilde{B}_n^+(s)}{\widetilde{W}_n(s)} ds - \int_0^{-\ln(1-a)} \frac{\lambda e^{-s}}{f_W(e^{-s})} ds| \\ \leqslant &|\int_0^{\widetilde{T}_n(a\lambda^{(n)}n)} \frac{\widetilde{B}_n^+(s)}{\widetilde{W}_n(s)} ds - \int_0^{\widetilde{T}_n(a\lambda^{(n)}n)} \frac{\lambda e^{-s}}{f_W(e^{-s})} ds| + |\int_{-\ln(1-a)}^{\widetilde{T}_n(a\lambda^{(n)}n)} \frac{\lambda e^{-s}}{f_W(e^{-s})} ds|.\end{aligned}$$

For n sufficiently large, we can find some (sufficiently large) constant C such that

$$-\ln(1 - a) - C < \widetilde{T}_n(a\lambda^{(n)}n) < -\ln(1 - a) + C.$$

Thus the first term converges to 0 in probability by Theorem 2.1 and the fact that $f_W(e^{-s}) \geqslant \epsilon$ on $[0, -\ln(1-a)+C]$ for some $\epsilon > 0$. Similarly, by Equation 2.31 and the boundedness of $\lambda e^{-s}/f_W(e^{-s})$ on the interval between $\widetilde{T}_n(a\lambda^{(n)}n)$ and $-\ln(1-a)$, the second term also converges to 0 in probability. Hence we have

$$T_n(a\lambda^{(n)}n) \xrightarrow{p} \int_0^{-\ln(1-a)} \frac{\lambda e^{-s}}{f_W(e^{-s})} ds.$$

Then through the change of variable, we have

$$\int_0^{-\ln(1-a)} \frac{\lambda e^{-s}}{f_W(e^{-s})} ds = \int_{1-a}^{1} \frac{\lambda}{f_W(e^{-t})} ds,$$

which completes the proof of Theorem 2.12.

2.8 Appendix

2.8.1 Proof of Lemmas

This appendix contains the proofs of all the lemmas in the main text. We start by recalling some classical results on death processes and martingale theory, on which relies the study of default cascade processes (as in [13, 152]).

Lemma 2.18. *(The Glivenko-Cantelli Theorem) Let $T_1, \dots, T_n$ be i.i.d random variables with distribution function $F(t) := \mathbb{P}(T_i \leqslant t)$. Let $X_n(t)$ be their empirical distribution function $X_n(t) := \#\{i \leqslant n : T_i \leqslant t\}/n$. Then $\sup_t |X_n(t) - F(t)| \xrightarrow{p} 0$ as $n \to \infty$.*

Since in the balls-and-bins model described in Section 2.2.3, every in ball dies independently after an exponential time with parameter 1, we have a pure death process starting with some number of balls whose lifetimes are i.i.d exp(1). As a corollary of the above lemma, we have:

Lemma 2.19. *(Death Process Lemma) Let $N^{(n)}(t)$ be the number of balls alive at time t and all balls have i.i.d. lifetime $\exp(1)$, starting with initial number $N^{(n)}(0) = n$. Then*

$$\sup_{t \geqslant 0} |N^{(n)}(t)/n - e^{-t}| \xrightarrow{p} 0 \qquad \text{as} \quad n \to \infty.$$

Our proof of the asymptotic normality for the default contagion is based on a martingale limit theorem in [147]. Let X be a martingale defined on $[0, \infty)$. We denote its quadratic variation by $[X, X]_t$. We also denote the (bilinear) covariation of two martingales X and Y by $[X, Y]_t$. In particular, if X and Y are two martingales with path-wise finite variation, then $[X, Y]_t := \sum_{0<s\leqslant t} \Delta X(s)\Delta Y(s)$, where $\Delta X(s) := X(s) - X(s-)$ is the jump of X at s and similarly $\Delta Y(s) := Y(s) - Y(s-)$. Note that in this chapter, the considered martingales are always RCLL (right continuous and with left limit) and have only finite number of jumps. Hence, the quadratic variation is finite. We also set $[X, Y]_0 = 0$. For two vector-valued martingales $\mathbf{X} = (X_i)_{i=1}^n$ and $\mathbf{Y} = (Y_j)_{j=1}^m$, we define $[\mathbf{X}, \mathbf{Y}]$ to be the $n \times m$ real matrix with every entry being $([\mathbf{X}, \mathbf{Y}])_{i,j} = [X_i, Y_j]$.

Theorem 2.20 (Martingale Limit Theorem [147]). *Assume that for each n, $\mathbf{M}^{(n)}(t) = (\mathbf{M}_i^{(n)}(t))_{i=1}^q$ is a q-dimensional martingale on $[0,\infty)$ with $\mathbf{M}^{(n)}(0) = \mathbf{0}$, and that $\Sigma(t)$ is a (nonrandom) continuous matrix-valued function satisfying, for every fixed $t \geqslant 0$,*

(i) $[\mathbf{M}^{(n)}, \mathbf{M}^{(n)}]_t \xrightarrow{p} \Sigma(t)$ *as* $n \to \infty$,

(ii) $\sup_n \mathbb{E}[M_i^{(n)}, M_i^{(n)}]_t < \infty$, *for all* $i = 1, \dots, q$.

Then $\mathbf{M}^{(n)} \xrightarrow{d} \mathbf{Z}$ *as* $n \to \infty$, *in* $\mathcal{D}[0,\infty)$, *where* $\mathbf{Z}$ *is a continuous q-dimensional Gaussian martingale with* $\mathbb{E}[\mathbf{Z}(t)] = \mathbf{0}$ *and covariances* $\mathbb{E}[\mathbf{Z}(t)\mathbf{Z}'(s)] = \Sigma(t)$, *for* $0 \leqslant t \leqslant s < \infty$.

This theorem yields joint convergence of the processes $\{M_i^{(n)}\}_{i=1}^q$ and can be extended to infinitely many processes (i.e., for the case $q = \infty$). Indeed, by definition, an infinite family of stochastic processes converge jointly if every finite subfamily does. We shall use the above theorem for stopped martingales.

2.8.2 Proof of Lemma 2.2

Recall that $f_W(z) := \lambda z - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=1}^{d_x^+} q_x(\theta)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big)$, and, $z^\star := \sup\big\{z \in [0,1] : f_W(z) = 0\big\}$. By the initial condition $q_x(0) > 0$ for some $x \in \mathcal{X}$, we have $f_W(1) > 0$ and $z^\star < 1$. Let us take a constant $t_1 > 0$ such that $t_1 < -\ln z^\star$. By continuity of $f_W(z)$ on $[0,1]$, it follows that $f_W(z) > 0$ on $(z^\star, 1]$. Hence, there exists some constant $C_1 > 0$ such that $f_W(e^{-t}) > C_1$ for $t \leqslant t_1$.

Since $W_n(\tau_n^\star) = -1$, if $\tau_n^\star \leqslant t_1$ then $W_n(\tau_n^\star)/n - f_W(e^{-\tau_n^\star}) \leqslant -C_1$ for n large. But on the other hand,

$$\sup_{t\leqslant\tau_n^\star} |\frac{W_n(t)}{n} - f_W(e^{-t}))| \xrightarrow{p} 0.$$

This is a contradiction. Therefore, we must have $\mathbb{P}(\tau_n^\star \leqslant t_1) \longrightarrow 0$, as $n \to \infty$. In the case $z^\star = 0$, we can take arbitrary t_1, which implies that $\tau_n^\star \xrightarrow{p} \infty$.

We now consider the case $z^\star \in (0,1]$. Let $\varepsilon > 0$ small enough and fix the constant $t_2 \in (-\ln z^\star, -\ln(z^\star - \varepsilon))$. By using a similar argument and given the assumption $f_W'(z^\star) > 0$, we can show there exists some constant $C_2 > 0$, such that $W_n(\tau_n^\star)/n - f_W(e^{-\tau_n^\star}) \geqslant C_2$ when n large if $\tau_n^\star \geqslant t_2$. Thus $\mathbb{P}(\tau_n \geqslant t_2) \longrightarrow 0$ as $n \to \infty$. Since t_1 and t_2 are arbitrary, tending both t_1 and t_2 to $-\ln z^\star$, implies that $\tau_n^\star \xrightarrow{p} -\ln z^\star$. This completes the proof of lemma.

2.8.3 Proof of Lemma 2.13

First note that $U_{x,\theta,0}^{(n)}(0) = N_{x,\theta}^{(n)}$ and $U_{x,\theta,s}^{(n)}(0) = 0$ for $\ell \geqslant 1$. Further, from Assumption 2.2, $N_x^{(n)}/n \to \mu_x$ and $q_x^{(n)}(\theta) \to q_x(\theta)$ as $n \to \infty$, for all $x \in \mathcal{X}$ and $\theta = 0, 1, \dots, d_x^+$. Moreover, by the strong law of large numbers, $N_{x,\theta}^{(n)}/N_x^{(n)} \to q_x^{(n)}(\theta)$ a.s. as $n \to \infty$.

Consider now the death process as described in Section 2.2.3. Let us fix $x \in \mathcal{X}$ and integers θ, r with $0 \leqslant r \leqslant \theta \leqslant d_x^+$. Consider the $N_{x,\theta}^{(n)}$ bins which starts with d_x^+ alive in balls. For $k = 1, \dots, N_{x,\theta}^{(n)}$, let T_k be the time that the $(d_x^+ - r)$-th ball is removed (killed) in the k-th such bin. Then we have $\#\{k : T_k \leqslant t\} = \sum_{s=0}^{r} U_{x,\theta,s}^{(n)}(t)$. Since the number of remaining balls in any of such bins at time t are i.i.d. random variables with distribution $\mathsf{Bin}(d_x^+, e^{-t})$, then we have $\mathbb{P}(T_k \leqslant t) = \sum_{s=0}^{r} b(d_x^+, e^{-t}, s)$. Hence, by using Glivenko-Cantelli theorem,

$$\sup_{t \leqslant \tau_n} \Big| \frac{1}{N_{x,\theta}^{(n)}} \sum_{s=r+1}^{d_x^+} U_{x,\theta,s}^{(n)}(t) - \sum_{s=r+1}^{d_x^+} b(d_x^+, e^{-t}, s) \Big| \xrightarrow{p} 0, \qquad \text{as} \quad N_{x,\theta}^{(n)} \to \infty.$$

Multiply the above equation by $N_{x,\theta}^{(n)}/N_x^{(n)}$, and by the law of large numbers, we have

$$\sup_{t \leqslant \tau_n} \Big| \frac{1}{N_x^{(n)}} \sum_{s=r+1}^{d_x^+} U_{x,\theta,s}^{(n)}(t) - q_x^{(n)}(\theta) \sum_{s=r+1}^{d_x^+} b(d_x^+, e^{-t}, s) \Big| \xrightarrow{p} 0, \qquad \text{as} \quad N_x^{(n)} \to \infty. \tag{2.32}$$

Moreover, by using Assumption 2.2,

$$\sup_{t \geqslant 0} \Big| \frac{N_x^{(n)}}{n} q_x^{(n)}(\theta) \sum_{s=r+1}^{d_x^+} b(d_x^+, e^{-t}, s) - \mu_x q_x(\theta) \sum_{s=r+1}^{d_x^+} b(d_x^+, e^{-t}, s) \Big| \longrightarrow 0, \quad \text{as} \quad n \to \infty.$$

By combining the two formulas above and multiplying (2.32) by $N_x^{(n)}/n$, we obtain

$$\sup_{t \leqslant \tau_n} \Big| \frac{1}{n} \sum_{s=r+1}^{d_x^+} U_{x,\theta,s}^{(n)}(t) - \mu_x q_x(\theta) \sum_{s=r+1}^{d_x^+} b(d_x^+, e^{-t}, s) \Big| \xrightarrow{p} 0, \quad \text{as} \quad n \to \infty. \tag{2.33}$$

Hence, by using (2.33) for $r_1 = \ell$ and $r_2 = \ell - 1$, and then taking the difference, we obtain

$$\sup_{t \leqslant \tau_n} \Big| \frac{U_{x,\theta,\ell}^{(n)}(t)}{n} - \mu_x q_x(\theta) b\left(d_x^+, e^{-t}, \ell\right) \Big| \xrightarrow{p} 0, \quad \text{as} \quad n \to \infty.$$

Note that the above equation holds for all $x \in \mathcal{X}$ and $\theta = 1, \dots, d_x^+$. Hence, the same convergence also holds for any partial sum over x and θ. In particular,

$$\sup_{t \leqslant \tau_n} \sum_{\theta=1}^{d_x^+} \sum_{s=d_x^+-\theta+1}^{d_x^+} \Big| \frac{U_{x,\theta,s}^{(n)}(t)}{n} - \mu_x q_x(\theta) b(d_x^+, e^{-t}, s) \Big| \xrightarrow{p} 0. \tag{2.34}$$

Let $\mathcal{X}_K$ be the set of all characteristic $x \in \mathcal{X}$ such that $d_x^+ + d_x^- \leqslant K$. Since (by Assumption 2.3a) $\lambda \in (0, \infty)$ then for arbitrary small $\varepsilon > 0$, there exists K_ε such that $\sum_{x \in \mathcal{X} \setminus \mathcal{X}_{K_\varepsilon}} \mu_x (d_x^+ + d_x^-) < \varepsilon$. Further, by Assumption 2.3a and dominated convergence,

$$\sum_{x \in \mathcal{X} \setminus \mathcal{X}_{K_\varepsilon}} (d_x^+ + d_x^-) N_x^{(n)}/n \to \sum_{x \in \mathcal{X} \setminus \mathcal{X}_{K_\varepsilon}} (d_x^+ + d_x^-) \mu_x < \varepsilon.$$

Hence for n large enough, we have $\sum_{x\in\mathcal{X}\setminus\mathcal{X}_{K_\varepsilon}}(d_x^+ + d_x^-)N_x^{(n)}/n < 2\varepsilon$. By (2.34), we obtain

$$\begin{aligned}
&\sup_{t\leqslant\tau_n}\sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+}\left|U_{x,\theta,s}^{(n)}(t)/n - \mu_x q_x(\theta)b(d_x^+, e^{-t}, s)\right| \\
\leqslant &\sup_{t\leqslant\tau_n}\sum_{x\in\mathcal{X}_{K_\varepsilon}}(d_x^+ + d_x^-)\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+}\left|U_{x,\theta,s}^{(n)}(t)/n - \mu_x q_x(\theta)b(d_x^+, e^{-t}, s)\right| \\
&+\sup_{t\leqslant\tau_n}\sum_{x\in\mathcal{X}\setminus\mathcal{X}_{K_\varepsilon}}(d_x^+ + d_x^-)\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+}\left|U_{x,\theta,s}^{(n)}(t)/n - \mu_x q_x(\theta)b(d_x^+, e^{-t}, s)\right| \\
\leqslant &o_p(1) + \sum_{x\in\mathcal{X}\setminus\mathcal{X}_{K_\varepsilon}}(d_x^+ + d_x^-)(N_x^{(n)}/n + \mu_x) \leqslant o_p(1) + 3\varepsilon.
\end{aligned}$$

Then let $\varepsilon\to 0$, it follows that (2.10) holds, which completes the proof of Lemma 2.13

2.8.4 Proof of Lemma 2.14

We proceed with the proof of asymptotic normality for $V_{x,\theta,s}^{(n)}$ similarly to the proof of [152, Theorem 3.1]. However, the proof will be more involved and includes more calculations. Since $V_{x,\theta,s}^{(n)}$ changes by -1 when one of the alive (in) balls in $U_{x,\theta,s}^{(n)}$ bins dies, and there are $sU_{x,\theta,s}^{(n)}(t)$ such balls at time t, we obtain

$$dV_{x,\theta,s}^{(n)}(t) = -sU_{x,\theta,s}^{(n)}(t)dt + d\mathcal{M}'_t,$$

where $\mathcal{M}'$ is a martingale.

We define further $\widehat{V}_{x,\theta,s}^{(n)}(t) := e^{st}V_{x,\theta,s}^{(n)}(t)$ and set the convention $V_{x,\theta,s}^{(n)}(t)\equiv 0$ for all $s > d_x^+$. Then by Ito's formula, for $0\leqslant s\leqslant d_x^+$

$$\begin{aligned}
d\widehat{V}_{x,\theta,s}^{(n)}(t) &= se^{st}V_{x,\theta,s}^{(n)}(t)dt + e^{st}dV_{x,\theta,s}^{(n)}(t) = se^{st}V_{x,\theta,s}^{(n)}(t)dt - se^{st}U_{x,\theta,s}^{(n)}(t)dt + e^{st}d\mathcal{M}'_t \\
&= se^{-t}\widehat{V}_{x,\theta,(s+1)}^{(n)}(t)dt + d\mathcal{M}_t,
\end{aligned}$$

where $d\mathcal{M}_t = e^{st}d\mathcal{M}'_t$ is also a martingale differential. Thus

$$M_{x,\theta,s}^{(n)}(t) := \widehat{V}_{x,\theta,s}^{(n)}(t) - s\int_0^t e^{-r}\widehat{V}_{x,\theta,(s+1)}^{(n)}(r)dr \tag{2.35}$$

is also a martingale for every $0\leqslant s\leqslant d_x^+$. We can calculate its quadratic variation by

$$\left[M_{x,\theta,s}^{(n)}, M_{x,\theta,s}^{(n)}\right]_t = \sum_{0<r\leqslant t}|\Delta M_{x,\theta,s}^{(n)}(r)|^2 = \sum_{0<r\leqslant t}|\Delta\widehat{V}_{x,\theta,s}^{(n)}(r)|^2 = \int_0^t e^{2sr}d(-V_{x,\theta,s}^{(n)}(r)). \tag{2.36}$$

Then,

$$\widetilde{M}_{x,\theta,s}^{(n)}(t) := n^{-1/2}\left(M_{x,\theta,s}^{(n)}(t) - M_{x,\theta,s}^{(n)}(0)\right),$$

is a martingale with initial value at 0. Denote $\varphi_{x,\theta,s}(y) := \mu_x q_x(\theta)\beta(d_x^+, y, s)$. Then by similar arguments as in the proof of [152, Theorem 3.1], we have

$$\widetilde{M}^{(n)}_{x,\theta,s}(t \wedge \tau_n) \xrightarrow{d} Y_{x,\theta,s}(t \wedge t_0) \qquad \text{in} \quad \mathcal{D}\,[0,\infty)\,, \tag{2.37}$$

where $Y_{x,\theta,s}$ is a continuous Gaussian process with $\mathbb{E}[Y_{x,\theta,s}(t)] = 0$ and covariance

$$\mathbb{E}\left[Y_{x,\theta,s}(t)Y_{x,\theta,s}(u)\right] = \int_0^t e^{2sr} d(-\varphi_{x,\theta,s}(e^{-r})), \qquad 0 \leqslant t \leqslant u < \infty.$$

Furthermore, for $s \neq r$, we can show that $V^{(n)}_{x,\theta,r}$ and $V^{(n)}_{x,\theta,s}$ never change simultaneously, almost surely. Thus, $[\widetilde{M}^{(n)}_{x,\theta,r}, \widetilde{M}^{(n)}_{x,\theta,s}]_t = 0$.

Hence, by Theorem 2.20 applied to the vector-valued martingale $\big(\widetilde{M}^{(n)}_{x,\theta,s}\big)_{s=0,\dots,d_x^+}$, the convergence holds jointly with a diagonal covariance matrix for $\big(Y_{x,\theta,s}\big)_{s=0,\dots,d_x^+}$, which implies that the processes $Y_{x,\theta,0}, \dots, Y_{x,\theta,d_x^+}$ are all independent.

As for two different types-thresholds (x,θ) and (x',θ'), the independence follows since for any $s = 0,\dots,d_x^+$ and $s' = 0,\dots,d_{x'}^+$, $V^{(n)}_{x,\theta,s}$ and $V^{(n)}_{x'\theta',s'}$ also a.s. never change simultaneously. This could also be observed from the nature of our balls and bins representation: the balls die independently and a.s. never die at the same moment. Hence, the death processes in bins with different types are independent.

We now compute $\widehat{V}^{(n)}_{x,\theta,s}$, using the Definition-Equation (2.35) for $M^{(n)}_{x,\theta,s}(t)$, repeatedly. We find that for $s = d_x^+$, $\widehat{V}^{(n)}_{x,\theta,d_x^+}(t) = M^{(n)}_{x,\theta,d_x^+}(t)$, and for $s = d_x^+ - 1$,

$$\widehat{V}^{(n)}_{x,\theta,s}(t) = M^{(n)}_{x,\theta,s}(t) + s\int_0^t e^{-r} M^{(n)}_{x,\theta,(s+1)}(r)dr.$$

Then for $s = d_x^+ - 2$, we obtain

$$\begin{aligned}
\widehat{V}^{(n)}_{x,\theta,s}(t) =& M^{(n)}_{x,\theta,s}(t) + s\int_0^t e^{-r} M^{(n)}_{x,\theta,(s+1)}(r)dr + \int_{r_2<r_1<t} s(s+1)e^{-(r_1+r_2)} M^{(n)}_{x,\theta,(s+2)}(r_2)dr_2dr_1 \\
=& M^{(n)}_{x,\theta,s}(t) + s\int_0^t e^{-r} M^{(n)}_{x,\theta,(s+1)}(r)dr + \int_0^t s(s+1)(e^{-r} - e^{-t})e^{-r} M^{(n)}_{x,\theta,(s+2)}(r)dr \\
=& M^{(n)}_{x,\theta,s}(t) + \sum_{j=s+1}^{d_x^+} s\binom{j-1}{s}\int_0^t (e^{-r} - e^{-t})^{j-s-1} e^{-r} M^{(n)}_{x,\theta,j}(r)dr.
\end{aligned}$$

Assume that the above formula holds for $s \leqslant k \leqslant d_x^+ - 1$. Then for $s-1$, we deduce

$$
\begin{aligned}
\widehat{V}^{(n)}_{x,\theta,s-1}(t) &= M^{(n)}_{x,\theta,s-1}(t) + s\int_0^t e^{-r}\widehat{V}^{(n)}_{x,\theta,s}(r)dr \\
&= M^{(n)}_{x,\theta,s-1}(t) + (s-1)\int_0^t e^{-r}M^{(n)}_{x,\theta,s}(r)dr + \sum_{j=s+1}^{d_x^+}\frac{s(s-1)}{j-s}\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s}e^{-r}M^{(n)}_{x,\theta,j}(r)dr \\
&= M^{(n)}_{x,\theta,s-1}(t) + \sum_{j=s}^{d_x^+}(s-1)\binom{j-1}{s-1}\int_0^t (e^{-r}-e^{-t})^{j-s}e^{-r}M^{(n)}_{x,\theta,j}(r)dr.
\end{aligned}
$$

By induction, we obtain that

$$
\widehat{V}^{(n)}_{x,\theta,s}(t) = M^{(n)}_{x,\theta,s}(t) + \sum_{j=s+1}^{d_x^+} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}M^{(n)}_{x,\theta,j}(r)dr.
$$

We next define $\widehat{v}^{(n)}_{x,\theta,s}(t)$ for all $t \geqslant 0$, as the conditional expectation of $\widehat{V}^{(n)}_{x,\theta,s}(t)$ given its initial value $V^{(n)}_{x,\theta,s}(0)$. Namely, we have

$$
\widehat{v}^{(n)}_{x,\theta,s}(t) := \mathbb{E}\big[\widehat{V}^{(n)}_{x,\theta,s}(t)|V^{(n)}_{x,\theta,s}(0)\big] = M^{(n)}_{x,\theta,s}(0) + \sum_{j=s+1}^{d_x^+} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}M^{(n)}_{x,\theta,j}(0)dr.
$$

Note that by definition, $\mathbb{E}\big[\widehat{V}^{(n)}_{x,\theta,s}(t)\big] = e^{st}\mathbb{E}[V^{(n)}_{x,\theta,s}(t)]$. Further, $V^{(n)}_{x,\theta,s}(t)$ is the number of bins with type x, threshold θ and with at least s balls at time t in the death process where balls die independently with rate 1 (without stopping). Then at time t, each of such bins has the binomial probability $\beta(d_x^+, e^{-t}, s)$ to have at least s alive balls remaining. The initial number is $V^{(n)}_{x,\theta,s}(0) = N^{(n)}_{x,\theta}$. Consequently for all $s = 0, \dots, d_x^+$, we have

$$
\widehat{v}^{(n)}_{x,\theta,s} = e^{st}N^{(n)}_{x,\theta}\beta(d_x^+, e^{-t}, s). \tag{2.38}
$$

We further define for $t \leqslant \tau_n^\star$,

$$
\widetilde{V}^{(n)}_{x,\theta,s}(t) := n^{-1/2}\big(\widehat{V}^{(n)}_{x,\theta,s}(t) - \widehat{v}^{(n)}_{x,\theta,s}\big). \tag{2.39}
$$

It is then clear that

$$
\widetilde{V}^{(n)}_{x,\theta,s}(t) = \widetilde{M}^{(n)}_{x,\theta,s}(t) + \sum_{j=s+1}^{d_x^+} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}\widetilde{M}^{(n)}_{x,\theta,j}(r)dr.
$$

We next apply Theorem 2.20 to the above finite sum and take the limit (in distribution) under the summation sign. It follows that

$$
\widetilde{V}^{(n)}_{x,\theta,s}(t\wedge\tau_n) \xrightarrow{d} \widetilde{\mathcal{Z}}_{x,\theta,s}(t\wedge t_0), \tag{2.40}
$$

in $\mathcal{D}[0,\infty)$, where

$$\widetilde{\mathcal{Z}}_{x,\theta,s}(t) := Y_{x,\theta,s}(t) + \sum_{j=s+1}^{d_x^+} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}Y_{x,\theta,j}(r)dr.$$

Note that, although the initial $V_{x,\theta,s}^{(n)}(0)$ is random, by the standard (multidimensional) central limit theorem applied to $N_x^{(n)} = n\mu_x^{(n)}$ i.i.d. random variables $\mathbb{1}\{\Theta_i^{(n)} = \theta\}$, we have

$$n^{-1/2}\Big(N_{x,\theta}^{(n)} - n\mu_x^{(n)}q_x^{(n)}(\theta)\Big) \xrightarrow{d} \mathcal{Y}_{x,\theta}^*, \quad \text{as} \quad n\to\infty, \tag{2.41}$$

where $\mathcal{Y}_{x,\theta}^* \sim \mathcal{N}(0,\mu_x q_x(\theta)(1-q_x(\theta)))$ and $\mathrm{Cov}(\mathcal{Y}_{x,\theta_1}^*,\mathcal{Y}_{x,\theta_2}^*) = -\mu_x q_x(\theta_1)q_x(\theta_2)$ for $\theta_1 \neq \theta_2$. We denote by (for all θ_1,θ_2)

$$\psi_{x,\theta_1,\theta_2} := \mathrm{Cov}(\mathcal{Y}_{x,\theta_1}^*,\mathcal{Y}_{x,\theta_2}^*).$$

Notice that for all $x_1 \neq x_2$, $\mathcal{Y}_{x_1,\theta_1}^*, \mathcal{Y}_{x_2,\theta_2}^*$ are independent and $\mathrm{Cov}(\mathcal{Y}_{x_1,\theta_1}^*,\mathcal{Y}_{x_2,\theta_2}^*) = 0$. Then we have by (2.38) and (2.41) that jointly for all triple (x,θ,s), in $\mathcal{D}[0,\infty)$ as $n\to\infty$,

$$n^{-1/2}\Big(\hat{v}_{x,\theta,s}^{(n)} - ne^{st}\mu_x^{(n)}q_x^{(n)}(\theta)\beta(d_x^+,e^{-t},s)\Big) \xrightarrow{d} \widehat{\mathcal{Z}}_{x,\theta,s}(t), \tag{2.42}$$

where $\widehat{\mathcal{Z}}_{x,\theta,s}(t)$ is also a Gaussian process with mean 0 and covariance

$$\mathbb{E}\big[\widehat{\mathcal{Z}}_{x,\theta,s}(t)\widehat{\mathcal{Z}}_{x,\theta,s}(u)\big] = e^{(t+u)s}\beta(d_x^+,e^{-t},s)\beta(d_x^+,e^{-u},s)\psi_{x,\theta,\theta}.$$

We now analyse the independence between $\widetilde{\mathcal{Z}}_{x,\theta,s}$ and $\widehat{\mathcal{Z}}_{x,\theta,s}$. Let combine x and Θ_x as a new type $\varrho := (x,\theta)$ in the set of all possible combination for $x\in\mathcal{X}$ and $\theta\in\mathbb{N}$. Notice that we have fixed proportion for the type distribution $\mu_x^{(n)}$, $x\in\mathcal{X}$ in the network. But the threshold Θ_x for all $x\in\mathcal{X}$ is random. Thus the proportion (denoted by $\mu_\varrho^{(n)}$) for the new type ϱ is random. It does not satisfy Assumption 2.3b, but satisfies in probability, i.e. replacing $O(n)$ by $O_p(n)$ in Assumption 2.3b. In addition, $\mu_\varrho^{(n)} \xrightarrow{p} \mu_\varrho$ with the limit distribution $\mu_\varrho := \mu_x q_x(\theta)$ for $\varrho = (x,\theta)$. Further, by using the Skorokhod's representation theorem [156, Theorem 3.30] as also stated in [152, Lemma 8.2], we can define all the processes on a new common probability space such that, for the type distribution, $\mu_\varrho^{(n)} \to \mu_\varrho$ and Assumption 2.3b hold almost surely.

Further, note that the distribution of $\widetilde{\mathcal{Z}}_{x,\theta,s}$ do not depend on the random proportion $\mu_\varrho^{(n)}$, but only on the limit distribution μ_ϱ. Hence the arguments in the above paragraph guarantee that conditioning on the initial value $V_{x,\theta,s}^{(n)}(0)$ does not have any influence on the distribution of $\widetilde{\mathcal{Z}}_{x,\theta,s}$. Therefore $\widetilde{\mathcal{Z}}_{x,\theta,s}$ and $\widehat{\mathcal{Z}}_{x,\theta,s}$ are independent for all (x,θ,s).

The above argument shows that $\widetilde{V}_{x,\theta,s}^{(n)}$ converges to a Gaussian process. We next define

$$\widetilde{\widetilde{V}}_{x,\theta,s}^{(n)}(t) := n^{-1/2}\Big(\widehat{V}_{x,\theta,s}^{(n)}(t) - ne^{st}\mu_x^{(n)}q_x^{(n)}(\theta)\beta(d_x^+,e^{-t},s)\Big).$$

By (2.39), (2.40) and (2.42), we obtain that $\widetilde{\widehat{V}}^{(n)}_{x,\theta,s}(t\wedge\tau_n) \xrightarrow{d} \widetilde{\widehat{\mathcal{Z}}}_{x,\theta,s}(t\wedge t_0)$, where $\widetilde{\widehat{\mathcal{Z}}}_{x,\theta,s} := \widehat{\mathcal{Z}}_{x,\theta,s}+\widetilde{\mathcal{Z}}_{x,\theta,s}$ is a Gaussian process with mean 0 and covariance

$$\mathbb{E}\big[\widetilde{\widehat{\mathcal{Z}}}_{x,\theta,s}(t)\widetilde{\widehat{\mathcal{Z}}}_{x,\theta,s}(u)\big] = \mathbb{E}\big[\widehat{\mathcal{Z}}_{x,\theta,s}(t)\widehat{\mathcal{Z}}_{x,\theta,s}(u)\big] + \mathbb{E}\big[\widetilde{\mathcal{Z}}_{x,\theta,s}(t)\widetilde{\mathcal{Z}}_{x,\theta,s}(u)\big],$$

for all $0 \leqslant t \leqslant u < \infty$.

Next, we define for all triple (x, θ, s),

$$V^{*(n)}_{x,\theta,s}(t) := e^{-st}\widetilde{\widehat{V}}^{(n)}_{x,\theta,s}(t), \quad \widetilde{\widetilde{\mathcal{Z}}}_{x,\theta,s}(t) := e^{-st}\widetilde{\mathcal{Z}}_{x,\theta,s}(t),$$

and

$$\widehat{\widehat{\mathcal{Z}}}_{x,\theta,s}(t) := e^{-st}\widehat{\mathcal{Z}}_{x,\theta,s}(t), \quad \mathcal{Z}^*_{x,\theta,s}(t) := e^{-st}\widetilde{\widehat{\mathcal{Z}}}_{x,\theta,s}(t).$$

Then we have

$$V^{*(n)}_{x,\theta,s}(t\wedge\tau_n) \xrightarrow{d} \mathcal{Z}^*_{x,\theta,s}(t\wedge t_0). \tag{2.43}$$

We define further

$$\widehat{\sigma}_{x,\theta_1,\theta_2,r,s}(y) := \mathrm{Cov}\big((\widehat{\widehat{\mathcal{Z}}}_{x,\theta_1,r}(-\ln y), \widehat{\widehat{\mathcal{Z}}}_{x,\theta_2,s}(-\ln y))\big),$$

and,

$$\widetilde{\sigma}_{x,\theta,r,s}(y) := \mathrm{Cov}\big(\widetilde{\widetilde{\mathcal{Z}}}_{x,\theta,r}(-\ln y), \widetilde{\widetilde{\mathcal{Z}}}_{x,\theta,s}(-\ln y)\big).$$

By using all the independence and covariance formulas above, it follows that

$$\begin{aligned}
\mathrm{Cov}\big(\mathcal{Z}^*_{x_1,\theta_1,s_1}(t), \mathcal{Z}^*_{x_2,\theta_2,s_2}(t)\big) =& 0, \quad \text{for all} \quad x_1 \neq x_2,\\
\mathrm{Cov}\big(\mathcal{Z}^*_{x,\theta_1,s_1}(t), \mathcal{Z}^*_{x,\theta_2,s_2}(t)\big) =& \widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(e^{-t}), \quad \text{for all} \quad \theta_1 \neq \theta_2,\\
\mathrm{Cov}\big(\mathcal{Z}^*_{x,\theta,s_1}(t), \mathcal{Z}^*_{x,\theta,s_2}(t)\big) =& \widehat{\sigma}_{x,\theta,\theta,s_1,s_2}(e^{-t}) + \widetilde{\sigma}_{x,\theta,s_1,s_2}(e^{-t}),
\end{aligned}$$

where

$$\begin{aligned}
&\widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(e^{-t}) = \beta(d_x^+, e^{-t}, s_1)\beta(d_x^+, e^{-t}, s_2)\psi_{x,\theta_1,\theta_2},\\
&\psi_{x,\theta,\theta} = \mu_x q_x(\theta)(1-q_x(\theta)), \quad \psi_{x,\theta_1,\theta_2} = -\mu_x q_x(\theta_1)q_x(\theta_2) \quad \text{for all} \quad \theta_1 \neq \theta_2.
\end{aligned} \tag{2.44}$$

Moreover, the covariance between $\mathcal{Z}^*_{x,\theta_1,s}(t)$ and $\mathcal{Y}^*_{x,\theta_2}$ is given by

$$\mathrm{Cov}\big(\mathcal{Z}^*_{x,\theta_1,s}(t), \mathcal{Y}^*_{x,\theta_2}\big) = \mathrm{Cov}\big(\beta(d_x^+, e^{-t}, s)\mathcal{Y}^*_{x,\theta_1}, \mathcal{Y}^*_{x,\theta_2}\big) = \beta(d_x^+, e^{-t}, s)\psi_{x,\theta_1,\theta_2},$$

and for $x_1 \neq x_2$, again by the independence the covariance is 0.

We now compute $\widetilde{\sigma}_{x,\theta,r,s}(y)$. Recall that

$$\mathrm{Cov}(Y_{x,\theta,s}(-\ln y), Y_{x,\theta,r}(-\ln y)) = \mathbb{1}\{r = s\}\int_y^1 u^{-2s}d\varphi_{x,\theta,s}(u).$$

Then we obtain

$$\begin{aligned}\mathrm{Var}\big(\widetilde{\mathcal{Z}}_{x,\theta,s}(-\ln y)\big) &= \int_y^1 v^{-2s} d\varphi_{x,\theta,s}(v) \\ &+ \sum_{j=s+1}^{d_x^+} s^2 \binom{j-1}{s}^2 \iint_{y<u<z<1} (u-y)^{j-s-1}(z-y)^{j-s-1}\Big(\int_z^1 v^{-2j} d\varphi_{x,\theta,j}(v)\Big) du dz \\ &= \frac{1}{2}\sum_{j=s}^{d_x^+} \binom{j-1}{s-1}^2 \int_y^1 (v-y)^{2j-2s} v^{-2j} d\varphi_{x,\theta,j}(v).\end{aligned}$$

For $r \geqslant s$, we can write $r = s+k$ for some $k \geqslant 1$, and deduce that

$$\mathrm{Cov}\big(\widetilde{\mathcal{Z}}_{x,\theta,s}(-\ln y), \widetilde{\mathcal{Z}}_{x,\theta,s+k}(-\ln y)\big) = \frac{1}{2}\sum_{j=s+k}^{d_x^+} \binom{j-1}{s-1}\binom{j-1}{s+k-1} \int_y^1 (v-x)^{2j-2s-k} v^{-2j} d\varphi_{x,\theta,j}(v).$$

Hence we have

$$\begin{aligned}\widetilde{\sigma}_{x,\theta,s,s+k}(y) :&= y^{2s+k}\mathrm{Cov}\big(\widetilde{\mathcal{Z}}_{x,\theta,s}(-\ln y), \widetilde{\mathcal{Z}}_{x,\theta,s+k}(-\ln y)\big) \\ &= \frac{1}{2} y^{2s+k} \sum_{j=s+k}^{d_x^+} \binom{j-1}{s-1}\binom{j-1}{s+k-1} \int_y^1 (v-y)^{2j-2s-k} v^{-2j} d\varphi_{x,\theta,j}(v).\end{aligned} \tag{2.45}$$

This completes the proof of Lemma 2.14.

2.8.5 Proof of Lemma 2.15

We only provide the proof for f_{H^+}. The proof for the case $\clubsuit \in \{S, D, I^+, W\}$ follows in the same way. For all $d, y \in \mathbb{N}$ and $z \in [0,1]$, let us consider the following function

$$h((d,y);z) := \sum_{\ell=d-y+1}^{d} \ell b(d,z,\ell).$$

We define a sequence of bi-dimensional nonnegative integer valued random variables $\{X_n\}$ and X with distributions $\mathbb{P}(X_n = (d,y)) = \sum_{x\in\mathcal{A}_d} \mu_x^{(n)} q_x^{(n)}(y)$, and $\mathbb{P}(X = (d,y)) = \sum_{x\in\mathcal{A}_d} \mu_x q_x(y)$, where $\mathcal{A}_d := \{x \in \mathcal{X} : d_x^+ = d\}$. Then it follows that $f_{H^+}^{(n)}(z) = \mathbb{E}h(X_n;z)$ and $f_{H^+}(z) = \mathbb{E}h(X;z)$. By Assumption 2.3a, we have $X_n \to X$ in distribution as $n \to \infty$. Moreover, for any $z \in [0,1]$, $0 \leqslant h((d,y);z) \leqslant d$. Thus, $h(X_n;z) \leqslant X_n^{(1)}$, where $X_n^{(1)}$ is first dimensional component of X_n. Note also that by Assumption 2.3b, $X_n^{(1)}$ is uniformly integrable. Hence we have (as $n \to \infty$) for all $z \in [0,1]$,

$$f_{H^+}^{(n)}(z) = \mathbb{E}h(X_n;z) \to \mathbb{E}h(X;z) = f_{H^+}(z).$$

Further, an elementary calculation gives that

$$\frac{\partial}{\partial z} b(d,z,\ell) = db(d-1,z,\ell-1) - db(d-1,z,\ell).$$

Combining now with the fact that $b(d,z,\ell) \in [0,1]$, we have $|\frac{\partial}{\partial z}b(d,z,\ell)| \leqslant d$. In addition, using a simple induction gives that for every $j \geqslant 0$, $|\frac{\partial^j}{\partial z^j}b(v,z,\ell)| \leqslant d^j$. We therefore obtain that

$$\Big|\frac{\partial^j}{\partial z^j}h(X_n;z)\Big| \leqslant (X_n^{(1)})^j \sum_{\ell=1}^{X_n^{(1)}} \ell \leqslant (X_n^{(1)})^{j+2}. \tag{2.46}$$

This is again, by Assumption 2.3b, uniformly integrable. Hence, we also have (as $n \to \infty$)

$$\frac{\partial^j}{\partial z^j}f_{H^+}^{(n)}(z) = \mathbb{E}\frac{\partial^j}{\partial z^j}h(X_n;z) \to \mathbb{E}\frac{\partial^j}{\partial z^j}h(X;z) = \frac{\partial^j}{\partial z^j}f_{H^+}(z)$$

for all $z \in [0,1]$. Moreover, (2.46) also implies all these derivatives are uniformly bounded. Thus by the Arzela-Ascoli theorem (see e.g., [156]), as $n \to \infty$, $f_{H^+}^{(n)}(z) \to f_{H^+}(z)$ together with all its derivatives uniformly on $[0,1]$. This completes the proof for $\clubsuit = H^+$ and the proof for the case $\clubsuit \in \{S, D, I^+, W\}$ follows in the same way.

2.8.6 Proof of Lemma 2.16

We use the same notations as in the proof of Lemma 2.14 and only provide the proof of (2.13). Other convergences are simpler and could be proved by using similar arguments. First note that, for $\widetilde{V}_{x,s}^{(n)}(t) = \sum_\theta \widetilde{V}_{x,\theta,s}^{(n)}(t)$, we have

$$\begin{aligned}\widetilde{V}_{x,s}^{(n)}(t) &= \sum_{\theta=1}^{d_x^+}\widetilde{M}_{x,\theta,s}^{(n)}(t) + \sum_{j=s+1}^{d_x^+} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}\sum_{\theta=1}^{d_x^+}\widetilde{M}_{x,\theta,j}^{(n)}(r)dr \\ &= \widetilde{M}_{x,s}^{(n)}(t) + \sum_{j=s+1}^{d_x^+} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}\widetilde{M}_{x,j}^{(n)}(r)dr,\end{aligned}$$

where $\widetilde{M}_{x,s}^{(n)}(t) := \sum_{\theta=1}^{d_x^+}\widetilde{M}_{x,\theta,s}^{(n)}(t)$. This is again a partial sum and Theorem 2.20 applies. We therefore have

$$\widetilde{V}_{x,s}^{(n)}(t\wedge\tau_n) \xrightarrow{d} \widetilde{\mathcal{Z}}_{x,s}(t\wedge t_0) := \sum_{\theta=1}^{d_x^+}\widetilde{\mathcal{Z}}_{x,\theta,s}(t\wedge t_0),$$

in $\mathcal{D}[0,\infty)$. More precisely,

$$\widetilde{M}_{x,s}^{(n)}(t\wedge\tau_n) \xrightarrow{d} Y_{x,s}(t\wedge t_0) \qquad \text{in} \quad \mathcal{D}[0,\infty),$$

where $Y_{x,s}$ is a continuous Gaussian process with $\mathbb{E}Y_{x,s}(t) = 0$ and covariance

$$\mathbb{E}\big[Y_{x,s}(t)Y_{x,s}(u)\big] = \sum_{\theta=1}^{d_x^+}\int_0^t e^{2sr}d(-\varphi_{x,\theta,s}(e^{-r})) = \int_0^t e^{2sr}d(-\varphi_{x,s}(e^{-r})),$$

for $0 \leqslant t \leqslant u < \infty$ and

$$\varphi_{x,s}(y) = \sum_{\theta=1}^{d_x^+} \varphi_{x,\theta,s}(y) = \mu_x \beta(d_x^+, y, s).$$

We now prove that the convergence also hold for an infinite sum which is used to prove our final result. Let

$$Q_{x,\theta,s}^{(n)}(t) := e^{-st} n^{-1/2} \big(\widehat{v}_{x,\theta,s}^{(n)} - n e^{st} \mu_x^{(n)} q_x^{(n)}(\theta) \beta(d_x^+, e^{-t}, s)\big). \tag{2.47}$$

Then we have for all (x, θ, s) and all $t > 0$,

$$\mathrm{Var}(Q_{x,\theta,s}^{(n)}(t)) \leqslant \mu_x^{(n)} q_x^{(n)}(\theta)(1 - q_x^{(n)}(\theta)).$$

Recall that $\mathcal{X}_s^+$ and $\mathcal{X}_s^-$ denote the set of characteristics which have in-degree $d_x^+ \geqslant s$ and out-degree $d_x^- \geqslant s$ respectively. Let Θ_x be an arbitrary subset of $\{1, \dots, d_x^+\}$. By (2.42), the convergence holds for the finite sum $\sum_{x \in \mathcal{X} \setminus \mathcal{X}_\ell^+} \sum_{\theta \in \Theta_x} Q_{x,\theta,s}^{(n)}(t)$. We now consider the following infinite sum

$$\sum_{x \in \mathcal{X}_\ell^+} (d_x^+ + d_x^-) \sum_{\theta \in \Theta_x} Q_{x,\theta,s}^{(n)}(t).$$

Since power function can be controlled by exponential function, there exists a constant $C > 1$ such that for all $t > 0$,

$$\begin{aligned}
\sum_{x \in \mathcal{X}_\ell^+} ((d_x^+)^2 + (d_x^-)^2) \mathrm{Var}(\sum_{\theta \in \Theta_x} Q_{x,\theta,s}^{(n)}(t)) &\leqslant \sum_{x \in \mathcal{X}_\ell^+} ((d_x^+)^2 + (d_x^-)^2) \sum_{\theta \in \Theta_x} \mathrm{Var}(Q_{x,\theta,s}^{(n)}(t)) \\
&\leqslant \sum_{x \in \mathcal{X}_\ell^-} (d_x^-)^2 \mu_x^{(n)} + \sum_{x \in \mathcal{X}_\ell^+} (d_x^+)^2 \mu_x^{(n)} \\
&\leqslant \sum_{x \in \mathcal{X}_\ell^-} C^{d_x^-} \mu_x^{(n)} + \sum_{x \in \mathcal{X}_\ell^+} C^{d_x^+} \mu_x^{(n)}.
\end{aligned}$$

Thus we have by Assumption 2.3b, for n large enough, as $\ell \to \infty$,

$$\mathbb{E}[\sup_{t>0} | \sum_{x \in \mathcal{X}_\ell^+} (d_x^+ + d_x^-) \sum_{\theta=1}^{d_x^+} Q_{x,\theta,s}^{(n)}(t)|] \to 0. \tag{2.48}$$

Then using the convergence of the partial sums of $Q_{x,\theta,s}^{(n)}$, we can extend the convergence to an infinite sum of $Q_{x,\theta,s}^{(n)}$, by using e.g. [63, Theorem 4.2]. Further, the limit is also continuous. By (2.47) and (2.42), we have in $\mathcal{D}[0, \infty)$,

$$\sum_{x \in \mathcal{X}} (d_x^+ + d_x^-) \sum_{\theta \in \Theta_x} Q_{x,\theta,s}^{(n)}(t) \xrightarrow{d} \sum_{x \in \mathcal{X}} (d_x^+ + d_x^-) \sum_{\theta \in \Theta_x} \widehat{\mathcal{Z}}_{x,\theta,s}(t). \tag{2.49}$$

On the other hand, for the following infinite sum, we have

$$\begin{aligned}
\bar{V}_s^{(n)}(t) &:= \sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x}\widetilde{V}_{x,\theta,s}^{(n)}(t)\\
&= \sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x}\widetilde{M}_{x,\theta,s}^{(n)}(t)\\
&\quad + \sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x}\sum_{j=s+1}^{d_x^+} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}\widetilde{M}_{x,\theta,j}^{(n)}(r)dr\\
&= \overline{M}_s^{(n)}(t) + \sum_{j=s+1}^{\infty} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}\overline{M}_j^{(n)}(r)dr,
\end{aligned}$$

where $\overline{M}_s^{(n)}(t) := \sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x}\widetilde{M}_{x,\theta,s}^{(n)}(t)$ is a martingale with initial value 0.

The quadratic variation of $\overline{M}_s^{(n)}$ is calculated as the following

$$[\overline{M}_s^{(n)}, \overline{M}_s^{(n)}]_t \leqslant 2e^{2st}\sum_{x\in\mathcal{X}_s^+}((d_x^+)^2 + (d_x^-)^2)V_{x,s}^{(n)}(0)/n.$$

Using Assumption 2.3b, for all $A > 1$, there exists constants C_s and $C_{s,A}$ such that

$$\begin{aligned}
\sum_{x\in\mathcal{X}_s^+}((d_x^+)^2 + (d_x^-)^2)V_{x,s}^{(n)}(0) &\leqslant \sum_{x\in\mathcal{X}_s^-}(d_x^-)^2V_{x,s}^{(n)}(0) + \sum_{x\in\mathcal{X}_s^+}(d_x^+)^2V_{x,s}^{(n)}(0)\\
&\leqslant A^{-s}\sum_{x\in\mathcal{X}}\mu_x^{(n)}((C_sA)^{d_x^+} + (C_sA)^{d_x^-}) \leqslant C_{s,A}A^{-s}n.
\end{aligned}$$

Thus for any $t > 0$ and a fixed T, by choosing $A = e^{2t+4T}$ we get

$$[\overline{M}_s^{(n)}, \overline{M}_s^{(n)}]_t \leqslant 2\sum_{x\in\mathcal{X}_s^+}((d_x^+)^2 + (d_x^-)^2)e^{2st}V_s^{(n)}(0)/n \leqslant C_{s,A}e^{-4Ts}.$$

By Doob's L^2 inequality, we have (for some constant $C'_{s,T}$)

$$\mathbb{E}[\sup_{t\leqslant T}(\overline{M}_s^{(n)}(t))^2] \leqslant 4\mathbb{E}[\overline{M}_s^{(n)}, \overline{M}_s^{(n)}]_T \leqslant C'_{s,T}e^{-4Ts}.$$

Then combining the Cauchy-Schwarz inequality, we obtain (for some constant $C''_{s,T}$)

$$\mathbb{E}[\sup_{t\leqslant T}|\overline{M}_s^{(n)}(t)|] \leqslant C''_{s,T}e^{-2Ts}. \tag{2.50}$$

Let us define

$$\xi_{N,n}(t) := \sum_{j=N}^{\infty} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}\overline{M}_j^{(n)}(r)dr.$$

Then by (2.50) and some (simple) calculations we find that

$$\mathbb{E}\big(\sup_{t\leqslant T}|\xi_{N,n}(t)|\big) \leqslant \sum_{j=N}^{\infty} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}\mathbb{E}\big[\sup_{t\leqslant T}|\overline{M}_j^{(n)}(t)|\big]dr$$
$$\leqslant C''_{s,T}T\sum_{j=N}^{\infty} s\binom{j-1}{s}(1-e^{-T})^{j-s-1}e^{-2Tj} \leqslant C''_{s,T}Tse^{sT}\sum_{j=N}^{\infty}e^{-2Tj},$$

which implies for any fixed s and T, $\mathbb{E}(\sup_{t\leqslant T}|\xi_{N,n}(t)|) \to 0$ as $N\to\infty$, uniformly in n.

Using again the convergence of the partial sums of $\widetilde{V}^{(n)}_{x,\theta,s}$, again by [63, Theorem 4.2], we can extend the convergence to some infinite sums of $\widetilde{V}^{(n)}_{x,\theta,s}$. It follows that in $\mathcal{D}[0,\infty)$,

$$\sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x} e^{-st}\widetilde{V}^{(n)}_{x,\theta,s}(t\wedge\tau_n) \xrightarrow{d} \sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x}\tilde{\tilde{\mathcal{Z}}}_{x,\theta,s}(t\wedge t_0). \tag{2.51}$$

Combining now (2.39), (2.42), (2.49) and (2.51), it then follows that jointly for any $s>0$,

$$\sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x} V^{*(n)}_{x,\theta,s}(t\wedge\tau_n) \xrightarrow{d} \sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x}\mathcal{Z}^*_{x,\theta,s}(t\wedge t_0),$$

and the partial sum also converges for any fixed r,

$$\sum_{s=1}^{r}\sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x} V^{*(n)}_{x,\theta,s}(t\wedge\tau_n) \xrightarrow{d} \sum_{s=1}^{r}\sum_{x\in\mathcal{X}}(d_x^+ + d_x^-)\sum_{\theta\in\Theta_x}\mathcal{Z}^*_{x,\theta,s}(t\wedge t_0). \tag{2.52}$$

Then notice that for any $x\in\mathcal{X}$, $d_x^+>0$, we have that

$$\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+2}^{d_x^+} V^{*(n)}_{x,\theta,s}(t) = \sum_{s=2}^{\infty}\sum_{x\in\mathcal{X}}\sum_{\theta\in\Theta_x} V^{*(n)}_{x,\theta,s}(t),$$

with taking $\Theta_x=\{1,\dots,d_x^+\}$. It remains to prove the convergence of the above infinite sum. Similarly, we define the following infinite tail sum

$$\bar{\xi}_{N,n}(t) := \sum_{s=N}^{\infty}\sum_{x\in\mathcal{X}}\sum_{\theta\in\Theta_x}\widetilde{V}^{(n)}_{x,\theta,s}.$$

Note that when s is large, $C''_{s,T}$ can be bounded by another constant C_T only depending on T. Then by the same way as above and (2.50), we obtain

$$\mathbb{E}\big(\sup_{t\leqslant T}|\bar{\xi}_{N,n}(t)|\big) \leqslant \sum_{s=N}^{\infty}\sum_{j=s+1}^{\infty} s\binom{j-1}{s}\int_0^t (e^{-r}-e^{-t})^{j-s-1}e^{-r}\mathbb{E}\big[\sup_{t\leqslant T}|\overline{M}_j^{(n)}(t)|\big]dr + \sum_{s=N}^{\infty}\overline{M}_s^{(n)}(t)$$
$$\leqslant C_T\sum_{s=N}^{\infty}e^{-2Ts} + \leqslant C_T T\sum_{j=N+1}^{\infty}\sum_{s=N}^{j-1} s\binom{j-1}{s}(1-e^{-T})^{j-s-1}e^{-2Tj}$$
$$\leqslant C_T\sum_{s=N}^{\infty}e^{-2Ts} + C_T\sum_{j=N+1}^{\infty} je^{jT}e^{-2Tj} \leqslant 2C_T\sum_{s=N}^{\infty}se^{-sT},$$

which implies that for any fixed $T > 0$, $\mathbb{E}(\sup_{t\leqslant T}|\xi_{N,n}(t)|) \to 0$ as $N \to \infty$, uniformly in n. Combine with (2.48), we therefore have that for any $T > 0$ fixed, as $N \to \infty$,

$$\mathbb{E}\Big[\sup_{t\leqslant T}\Big|\sum_{x\in\mathcal{X}_\ell^+}\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+2}^{d_x^+} V_{x,\theta,s}^{*(n)}(t)\Big|\Big]\to 0,$$

Hence the same argument [63, Theorem 4.2] allows us to pass the limit under the infinite sum and with the limit being continuous. It then follows that, by using (2.52) and letting $r \to \infty$, to obtain that in $\mathcal{D}[0,\infty)$,

$$\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+2}^{d_x^+} V_{x,\theta,s}^{*(n)}(t\wedge\tau_n) \xrightarrow{d} \sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+2}^{d_x^+} \mathcal{Z}_{x,\theta,s}^{*}(t\wedge t_0).$$

2.8.7 Covariances in Theorem 2.7

Using the notation of Section 2.7.3 and 2.7.4, we now compute the covariances in Theorem 2.7 for the continuous Gaussian processes $\mathcal{Z}_S, \mathcal{Z}_{H^+}, \mathcal{Z}_{I^+}$ and $\mathcal{Z}_W$. For convenience, we make a change of variable $y = e^{-t}$, which decreases from 1 to 0 as t varies from 0 to ∞. We use the notation

$$\sigma_L(y) := \mathrm{Var}(\mathcal{Z}_L(-\ln y)), \quad \sigma_{x,\theta,s}^L(y) := \mathrm{Cov}\big(\mathcal{Z}_L(-\ln y), \mathcal{Z}_{x,\theta,s}^*(-\ln y)\big).$$

In order to compute $\sigma_{x,\theta,s}^L$, we apply Theorem 2.20 to $\widetilde{L}_n$ and $\widetilde{M}_{x,\theta,s}^{(n)}$ for all $s = d_x^+ - \theta + 1, \ldots, d_x^+$. Observe that each time $V_{x,\theta,s}^{(n)}$ decreases by 1, also an in ball dies and thus L_n decreases by 1. Hence, the quadratic covariation is

$$\begin{aligned}
[\widetilde{M}_{x,\theta,s}^{(n)}, \widetilde{L}_n]_t &= n^{-1}\sum_{0<r\leqslant t}\Delta M_{x,\theta,s}^{(n)}(r)\Delta\widehat{L}_n(r) = n^{-1}\sum_{0<r\leqslant t}\Delta\widehat{V}_{x,\theta,s}^{(n)}(r)\Delta\widehat{L}_n(r) \\
&= n^{-1}\sum_{0<r\leqslant t} e^{(s+1)r}\Delta V_{x,\theta,s}^{(n)}(r)\Delta L_n(r) = n^{-1}\int_0^t e^{(s+1)r} d(-V_{x,\theta,s}^{(n)}(r)).
\end{aligned}$$

Using integration by parts as before, we obtain

$$[\widetilde{M}_{x,\theta,s}^{(n)}, \widetilde{L}_n]_t = \int_0^t e^{(s+1)r} d(-\varphi_{x,\theta,s}(e^{-r})) + o_p(1) = \int_{e^{-t}}^1 u^{-(s+1)} d\varphi_{x,\theta,s}(u) + o_p(1).$$

We can then compute all needed covariances. First, the above analysis together with Theorem 2.20, gives that for all (x,θ,s),

$$\mathrm{Cov}\big(Y_{x,\theta,s}(-\ln y), \widetilde{\mathcal{Z}}_L(-\ln y)\big) = \int_y^1 u^{-(s+1)} d\varphi_{x,\theta,s}(u),$$

where $Y_{x,\theta,s}$ is defined in (2.37). For $\sigma_L(y)$ we have

$$\sigma_L(y) := \mathrm{Var}(\mathcal{Z}_L(-\ln y)) = \mathrm{Var}(y\widetilde{\mathcal{Z}}_L(-\ln y)) = \lambda(y - y^2)/2. \tag{2.53}$$

The above analysis together with Theorem 2.20, give that for all x, θ,

$$\operatorname{Cov}\big(Y_{x,\theta,s}(-\ln y), \widetilde{\mathcal{Z}}_L(-\ln y)\big) = \int_y^1 u^{-(s+1)} d\varphi_{x,\theta,s}(u).$$

On the other hand, for $v \leqslant t$, $\operatorname{Cov}(Y_{x,\theta,s}(v), \widetilde{\mathcal{Z}}_L(t)) = \operatorname{Cov}(Y_{x,\theta,s}(v), \widetilde{\mathcal{Z}}_L(v))$. Thus we have

$$\begin{aligned}\operatorname{Cov}(\widetilde{\mathcal{Z}}_{x,\theta,s}(-\ln y), \widetilde{\mathcal{Z}}_L(-\ln y)) =& \operatorname{Cov}(Y_{x,\theta,s}(t), \widetilde{\mathcal{Z}}_L(t)) + \int_0^t (e^{-r} - e^{-t})^{j-s-1} e^{-r} \operatorname{Cov}(Y_{x,\theta,s}(r), \widetilde{\mathcal{Z}}_L(r)) dr \\ =& \int_y^1 u^{-(s+1)} d\varphi_{x,\theta,s}(u) + \sum_{j=s+1}^{d_x^+} s \binom{j-1}{s} f_{sj}(y),\end{aligned}$$

where, with a change of variable $u = e^{-r}$, the function $f_{sj}(y)$ is defined as

$$\begin{aligned}f_{sj}(y) :&= \int_y^1 (u-y)^{j-s-1} \int_u^1 v^{-(j+1)} d\varphi_{x,\theta,j}(v) du = \int_y^1 \int_y^v (u-y)^{j-s-1} v^{-(j+1)} du d\varphi_{x,\theta,j}(v) \\ &= \frac{1}{j-s} \int_y^1 (v-y)^{j-s} v^{-(j+1)} d\varphi_{x,\theta,j}(v).\end{aligned}$$

We thus obtain

$$\operatorname{Cov}(\widetilde{\mathcal{Z}}_{x,\theta,s}(-\ln y), \widetilde{\mathcal{Z}}_L(-\ln y)) = \sum_{j=s}^{d_x^+} \binom{j-1}{s-1} \int_y^1 (v-y)^{j-s} v^{-(j+1)} d\varphi_{x,\theta,j}(v).$$

Also, $\operatorname{Cov}(\widehat{\mathcal{Z}}_{x,\theta,s}(-\ln y), \widetilde{\mathcal{Z}}_L(-\ln y)) = 0$ since they are independent. Then we conclude that

$$\begin{aligned}\sigma^L_{x,\theta,s}(y) &= y^{s+1} \operatorname{Cov}(\widetilde{\mathcal{Z}}_{x,\theta,s}(-\ln y), \widetilde{\mathcal{Z}}_L(-\ln y)) \\ &= y^{s+1} \sum_{j=s}^{d_x^+} \binom{j-1}{s-1} \int_y^1 (v-y)^{j-s} v^{-(j+1)} d\varphi_{x,\theta,j}(v).\end{aligned} \tag{2.54}$$

We can write now the covariances for the processes $\mathcal{Z}_S$, $\mathcal{Z}_{H^+}$, $\mathcal{Z}_{I^+}$ and $\mathcal{Z}_W$ by using $\sigma_L(y)$ and $\sigma^L_{x,\theta,s}(y)$ (computed above), $\widetilde{\sigma}_{x,\theta,r,s}(y)$ and $\widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}$ (given in Lemma 2.14). We only write the covariances between $\mathcal{Z}_S$, $\mathcal{Z}_{H^+}$ and $\mathcal{Z}_W$; the covariances of $\mathcal{Z}_{I^+}$ can easily be deduced from those of $\mathcal{Z}_{H^+}$. For convenience, we set $\pi_x(\theta) := d_x^+ - \theta + 1$. For the variances, by using (2.17)-(2.20), we have:

$$\begin{aligned}\sigma_{W,W}(y) = \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \Big[(d_x^-)^2 \widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) - 2 d_x^- \sigma^L_{x,\theta,\pi_x(\theta)}(y) \Big] + \sigma_L(y) \\ + \sum_{x\in\mathcal{X}} (d_x^-)^2 \sum_{\theta_1=1}^{d_x^+} \sum_{\theta_2=1}^{d_x^+} \widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y),\end{aligned} \tag{2.55}$$

$$\sigma_{S,S}(y) = \sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) + \sum_{x\in\mathcal{X}}\sum_{\theta_1=1}^{d_x^+}\sum_{\theta_2=1}^{d_x^+}\widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y), \tag{2.56}$$

and,

$$\begin{aligned}
\sigma_{H^+,H^+}(y) = &\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\Big[\pi_x^2(\theta)\widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) + 2\sum_{s=\pi_x(\theta)+1}^{d_x^+}\pi_x(\theta)\widetilde{\sigma}_{x,\theta,\pi_x(\theta),s}(y)\\
&+ \sum_{s=\pi_x(\theta)+1}^{d_x^+}\sum_{r=\pi_x(\theta)+1}^{d_x^+}\widetilde{\sigma}_{x,\theta,r,s}(y)\Big]\\
&+\sum_{x\in\mathcal{X}}\sum_{\theta_1=1}^{d_x^+}\sum_{\theta_2=1}^{d_x^+}\Big[\sum_{s_1=\pi_x(\theta_1)+1}^{d_x^+}\sum_{s_2=\pi_x(\theta_2)+1}^{d_x^+}\pi_x(\theta_1)\pi_x(\theta_2)\widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y)\\
&+\widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(y) + 2\pi_x(\theta_2)\sum_{s_2=\pi_x(\theta_2)+1}^{d_x^+}\widehat{\sigma}_{x,\theta_1,\theta_2,s,\pi_x(\theta_2)}(y)\Big].
\end{aligned} \tag{2.57}$$

For the covariances, by using again (2.17)-(2.20), we have:

$$\begin{aligned}
\sigma_{S,H^+}(y) = &\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\Big[\pi_x(\theta)\widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) + \sum_{s=\pi_x(\theta)+1}^{d_x^+}\widetilde{\sigma}_{x,\theta,\pi_x(\theta),s}(y)\Big]\\
&+\sum_{x\in\mathcal{X}}\sum_{\theta_1,\theta_2=1}^{d_x^+}\Big[\pi_x(\theta_1)\widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y) + \sum_{s=\pi_x(\theta_1)+1}^{d_x^+}\widehat{\sigma}_{x,\theta_1,\theta_2,s,\pi_x(\theta_2)}(y)\Big],
\end{aligned} \tag{2.58}$$

$$\sigma_{S,W}(y) = \sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\Big[\sigma^L_{x,\theta,\pi_x(\theta)}(y) - d_x^-\widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y)\Big] - \sum_{x\in\mathcal{X}}d_x^-\sum_{\theta_1=1}^{d_x^+}\sum_{\theta_2=1}^{d_x^+}\widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y), \tag{2.59}$$

and,

$$\begin{aligned}
\sigma_{H^+,W}(y) = &\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\Big[-d_x^-\pi_x(\theta)\widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) + \pi_x(\theta)\sigma^L_{x,\theta,\pi_x(\theta)}(y)\\
&+\sum_{s=\pi_x(\theta)+1}^{d_x^+}\big(\sigma^L_{x,\theta,s}(y) - d_x^-\widetilde{\sigma}_{x,\theta,\pi_x(\theta),s}(y)\big)\Big]\\
&-\sum_{x\in\mathcal{X}}d_x^-\sum_{\theta_1,\theta_2=1}^{d_x^+}\Big[\sum_{s=\pi_x(\theta_1)+1}^{d_x^+}\widehat{\sigma}_{x,\theta_1,\theta_2,s,\pi_x(\theta_2)}(y) + \pi_x(\theta_1)\widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y)\Big].
\end{aligned} \tag{2.60}$$

As for the covariances of $\mathcal{Z}_{I^+}$, we can deduce them from the above formulas by using (2.19).

2.8.8 Covariances in Theorem 2.9

We provide below the variances for the Gaussian processes $\mathcal{Z}_\Diamond(t)$ and $\mathcal{Z}^\star_\Diamond$ in Theorem 2.9. First, let us define

$$\mathcal{Z}^{(1)}_\Diamond := \sum_{x\in\mathcal{X}} \bar{L}^\odot_x \sum_{\theta=1}^{d_x^+} \mathcal{Z}^*_{x,\theta,d_x^+-\theta+1}(t),$$

$$\mathcal{Z}^{(2)}_\Diamond := \sum_{x\in\mathcal{X}} \bar{L}^\Diamond_x \sum_{\theta=1}^{d_x^+} [(\theta-1)\mathcal{Z}^*_{x,\theta,d_x^+-\theta+1}(t) - \sum_{s=d_x^+-\theta+1+1}^{d_x^+} \mathcal{Z}^*_{x,\theta,s}(t)].$$

Let $\sigma_{i,j}(e^{-t}) := \mathrm{Cov}(\mathcal{Z}^{(i)}_\Diamond(t), \mathcal{Z}^{(j)}_\Diamond(t))$, $i,j=1,2$. The variance $\sigma_\Diamond(t)$ is therefore

$$\sigma_\Diamond(t) = \sigma_{1,1}(e^{-t}) + 2\sigma_{1,2}(e^{-t}) + \sigma_{2,2}(e^{-t}), \tag{2.61}$$

where $\sigma_{i,j}(y)$ (for $i,j=1,2$) are calculated in the following. By using results of Section 2.8.7, we have (recall that $\pi_x(\theta) := d_x^+ - \theta + 1$):

$$\sigma_{1,1}(y) = \sum_{x\in\mathcal{X}} (\bar{L}^\odot_x)^2 \Big(\sum_{\theta=1}^{d_x^+} \widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) + \sum_{\theta_1,\theta_2=1}^{d_x^+} \widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y) \Big),$$

$$\begin{aligned}\sigma_{2,2}(y) = &\sum_{x\in\mathcal{X}} (\bar{L}^\Diamond_x)^2 \sum_{\theta=1}^{d_x^+} \Big[(\theta-1)^2 \widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) - 2(\theta-1) \sum_{s=\pi_x(\theta)+1}^{d_x^+} \widetilde{\sigma}_{x,\theta,\pi_x(\theta),s}(y) \\ &+ \sum_{r,s=\pi_x(\theta)+1}^{d_x^+} \widetilde{\sigma}_{x,\theta,r,s}(y) \Big] + \sum_{x\in\mathcal{X}} (\bar{L}^\Diamond_x)^2 \sum_{\theta_1,\theta_2=1}^{d_x^+} \Big[\sum_{s_1=\pi_x(\theta_1)+1}^{d_x^+} \sum_{s_2=\pi_x(\theta_2)+1}^{d_x^+} \widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(y) \\ &+ (\theta_1-1)(\theta_2-1)\widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y) - 2(\theta_2-1) \sum_{s=\pi_x(\theta_1)+1}^{d_x^+} \widehat{\sigma}_{x,\theta_1,\theta_2,s,\pi_x(\theta_2)}(y) \Big],\end{aligned}$$

and,

$$\begin{aligned}\sigma_{1,2}(y) = &\sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \Big[(\theta-1)\widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) - \sum_{s=\pi_x(\theta)+1}^{d_x^+} \widetilde{\sigma}_{x,\theta,\pi_x(\theta),s}(y) \Big] \\ &+ \sum_{x\in\mathcal{X}} \sum_{\theta_1,\theta_2=1}^{d_x^+} \Big[(\theta_1-1)\widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y) - \sum_{s=\pi_x(\theta_1)+1}^{d_x^+} \widehat{\sigma}_{x,\theta_1,\theta_2,s,\pi_x(\theta_2)}(y) \Big].\end{aligned}$$

Now we compute the variance of $\mathcal{Z}^\star_\Diamond$. Since it has been shown that

$$\mathcal{Z}^\star_\Diamond := \mathcal{Z}_\Diamond(t^\star) - \alpha^{-1} f'_\Diamond(z^\star)\mathcal{Z}_W(t^\star) = \mathcal{Z}_\Diamond(t^\star) - \Delta(z^\star)\mathcal{Z}_W(t^\star),$$

with $\Delta(z^\star) := \alpha^{-1} f'_\Diamond(z^\star)$, we have that

$$\sigma^\star_\Diamond := \sigma_\Diamond(t^\star) + \Delta(z^\star)^2 \sigma_{W,W}(z^\star) - 2\Delta(z^\star)\sigma_{\Diamond,W}(z^\star). \tag{2.62}$$

We computed $\sigma_\Diamond$ and $\sigma_{W,W}$ in Appendix 2.8.7. We have

$$\sigma_{\Diamond,W}(e^{-t}) = -\mathrm{Cov}(Z_\Diamond^{(1)}(t), Z_W(t)) - \mathrm{Cov}(Z_\Diamond^{(2)}(t), Z_W(t)) = -\sigma_{\Diamond,W}^{(1)}(e^{-t}) - \sigma_{\Diamond,W}^{(2)}(e^{-t}),$$

where

$$\begin{aligned}\sigma_{\Diamond,W}^{(1)}(y) &= \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} \sum_{\theta=1}^{d_x^+} \widehat{\sigma}_{x,\theta,\pi_x(\theta)}(y) - \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} d_x^- \sum_{\theta=1}^{d_x^+} \widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) \\ &\quad - \sum_{x\in\mathcal{X}} \bar{L}_x^{\odot} d_x^- \sum_{\theta_1,\theta_2=1}^{d_x^+} \widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y),\end{aligned}$$

and,

$$\begin{aligned}\sigma_{\Diamond,W}^{(2)}(y) &= \sum_{x\in\mathcal{X}} \bar{L}_x^{\Diamond} \sum_{\theta=1}^{d_x^+} \Big[-d_x^-(\theta-1)\widetilde{\sigma}_{x,\theta,\pi_x(\theta),\pi_x(\theta)}(y) + (\theta-1)\widehat{\sigma}_{x,\theta,\pi_x(\theta)}(y) \\ &\qquad - \sum_{s=\pi_x(\theta)+1}^{d_x^+} \Big(\sigma_{x,\theta,s}^L(y) - d_x^- \widetilde{\sigma}_{x,\theta,\pi_x(\theta),s}(y) \Big) \Big] \\ &\quad + \sum_{x\in\mathcal{X}} \bar{L}_x^{\Diamond} d_x^- \sum_{\theta_1,\theta_2=1}^{d_x^+} \Big[\sum_{s=\pi_x(\theta_1)+1}^{d_x^+} \widehat{\sigma}_{x,\theta_1,\theta_2,s,\pi_x(\theta_2)}(y) - (\theta_1-1)\widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y) \Big].\end{aligned}$$

2.9 Concluding Remarks

In this chapter, we propose a general tractable framework to study default cascades and systemic risk in a heterogeneous financial network, subject to an exogenous macroeconomic shock. We state various limit theorems for the final state of default contagion and systemic risk depending on the network structure and institutions' (observable) characteristics.

Our central limit theorems can be used to provide confidence intervals for the final fraction of defaults and systemic risk. As Figure 2.2 shows, the asymptotic normality turns out to be quite reliable for not necessarily very large network size. Our asymptotic results could also be made of great use in a more complex contagion model including fire sales [16].

The closed form interpretable limit theorems that we provide could also serve as a mandate for regulators to collect data on those specific network characteristics and assess systemic risk via more intensive computational methods.

It would be interesting to extend the optimal interventions model of Section 2.5 to a continuous-time Markov decision process by the planner, when the links (starting from fundamental defaults) are revealed one by one, and consequently, the planner can decide at any time to intervene or not. This would lead to a Markov decision problem and one could solve it (under some regularity assumptions) by using a dynamic programming approach. We refer to [26] for a similar model in a simpler setup with a core-periphery network structure.

Our results can also be used in a regulatory risk management framework, when a regulator imposes capital requirements on each bank. In practice, the capital ratio constraint is the same for all banks. However, using our heterogeneous setup, we could allow the regulator to choose optimally this capital ratio according to the type of the banks. The regulator's problem is then to choose the minimum capital ratio for each institution (according to it's type) so that the systemic risk (e.g., expected shortfall of external wealth under some random shocks applied to capitals) is below a certain critical value. We leave this and some related issues to a future work.

Chapter 3

Fire Sales and Default Cascades

This chapter is based on paper [2] in the publication list of Section 1.5.

Abstract. We present a general tractable framework for understanding the joint impact of fire sales and default cascades on systemic risk in complex financial networks. Our limit theorems quantify how price-mediated contagion across institutions with common asset holdings can worsen cascades of insolvencies in a heterogeneous financial network during a financial crisis. For given prices of illiquid assets, we show that, under some regularity assumptions, the default cascade model can be transferred to a death process problem. We model the price impact using a specified inverse demand function. Various limit theorems concerning the total shares sold and the equilibrium price of illiquid assets in a stylized fire sales model are stated. In particular, we show that the equilibrium prices of illiquid assets have asymptotically Gaussian fluctuations. Our numerical studies investigate the effect of heterogeneity in network structure and price impact function on the final size of the default cascade and fire sales loss.

Keywords: Fire Sales, Default Contagion, Financial Networks, Systemic Risk.

Contents

2.1 Introduction **36**
2.2 Model **39**
2.2.1 Financial Network and Default Cascade 39
2.2.2 Node Classification and Configuration Model 40
2.2.3 Death Process and Final Solvent Institutions 41
2.3 Limit Theorems **43**
2.3.1 Asymptotic Magnitude of Default Contagion 44
2.3.2 Asymptotic Normality of Default Contagion 46
2.4 Quantifying Systemic Risk **49**
2.4.1 Numerical analysis on heterogeneity and stability 51

2.5 Targeting Interventions in Financial Networks . . . 52
2.6 Default cascade process . . . 54
2.7 Proof of Main Theorems . . . 56
2.7.1 Proof of Theorem 2.1 . . . 56
2.7.2 Proof of Theorem 2.4 . . . 57
2.7.3 Proof of Theorem 2.6 . . . 58
2.7.4 Proof of Theorem 2.7 . . . 59
2.7.5 Proof of Theorem 2.8 . . . 62
2.7.6 Proof of Theorem 2.9 . . . 63
2.7.7 Proof of Theorem 2.10 . . . 64
2.7.8 Proofs in §2.6 . . . 65
2.8 Appendix . . . 70
2.8.1 Proof of Lemmas . . . 70
2.8.2 Proof of Lemma 2.2 . . . 71
2.8.3 Proof of Lemma 2.13 . . . 71
2.8.4 Proof of Lemma 2.14 . . . 73
2.8.5 Proof of Lemma 2.15 . . . 78
2.8.6 Proof of Lemma 2.16 . . . 79
2.8.7 Covariances in Theorem 2.7 . . . 83
2.8.8 Covariances in Theorem 2.9 . . . 86
2.9 Concluding Remarks . . . 87

3.1 Introduction

Financial institutions are interconnected in various ways. The global financial crisis of 2007-2009 simultaneously highlighted the importance of interbank network structure and fire sales in the amplification and transmission of initial shocks across the wider financial system.

This current chapter examines the combined impact of fire sales and default cascades on systemic risk in complex financial networks during a financial crisis. Fire sales refer to situations in which an institution attempts or is forced to sell a large volume of assets within a short period of time.

We consider a financial network in which institutions hold interbank liabilities, cash, and shares of one or multiple illiquid assets. When a firm defaults, its counterparties may sell their illiquid assets (deleveraging) in response to the losses they face due to this default, potentially triggering lower prices for these or related assets. This may lead to contagion of losses across institutions with common asset holdings. Indeed, marking to market of institutions' balance sheets reinforces network contagion: lower asset prices may force other institutions to default on their interbank liabilities. This results in an entanglement of price-mediated contagion and interbank network-mediated contagion.

We consider a random graph approach, which is appropriate for dealing with systemic risk in financial networks when only partial information on linkages is available, as pointed out in, for example, [30, 76, 126, 187]. We reduce the dimension of the problem by classifying financial institutions into different categories. This can be calibrated to real-world data using machine learning techniques for classification. Due to its tractability and interpretability, as well as its potential to be enriched with clustering (see, for example, [98, 195]), we use the configuration model as our base probabilistic model. The configuration model has been previously used to model the pure default cascade process in financial networks, as seen in [20, 27, 28].

We present a general tractable framework for understanding the joint impact of fire sales and default cascades on systemic risk in a heterogeneous financial network, subject to an exogenous macroeconomic shock. As demonstrated in Chapter 2, under some regularity assumptions on the network, the pure default cascade model can be transformed into a death process problem. Since our model is static in nature, following [22, 92, 133], we assume that all the liquidations occur simultaneously and instantaneously. We model the price impact using a given inverse demand function.

We state various limit theorems concerning the total sold shares and the equilibrium price of illiquid assets in a stylized fire sales model. In particular, we demonstrate that the equilibrium prices of illiquid assets exhibit asymptotically Gaussian fluctuations. Our numerical studies explore the effect of heterogeneity in network structure and price impact function on the final size of the default cascade and fire sales loss.

Literature review. The literature on financial networks and systemic risk is vast; see e.g., [91, 146] for surveys and references therein. Much research in this area focuses on an equilibrium approach to derive recovery rates from some fixed-point equations, as seen in e.g., [111, 114, 185]. This relies on the assumption that all debts are instantaneously cleared, which is unlikely to hold during a financial crisis. Many studies incorporate various channels through which risk spreads in financial networks within the framework introduced by Eisenberg and Noe [111]. To name just a few, [131, 185] introduce bankruptcy costs and mark-to-market losses. Following [20, 24, 93], we consider in this chapter recovery rates as given. The model can be extended to a setup with random recovery rates satisfying some cash-flow consistency conditions as discussed in [24].

Our work is related to the literature on the impact of network structure and heterogeneity on default contagion and systemic risk; see e.g., [1, 7, 25, 46, 77, 87, 105, 124, 126, 132, 179, 196]. In particular, [1] compares the ring and complete network structures, finding that the completely connected system is the most stable for small shocks but the least stable for large shocks (and vice-versa for the ring network). In [21], the authors present a more general framework to find the optimal network structure for reducing systemic risk and show that the optimal network compression problem is generically NP-hard. Our work is also related to the literature on central limit theorems for credit contagion and portfolio losses, see e.g., [129, 130]. The economics of contagious phenomena with heterogeneous agents goes back to [121].

Price-mediated contagion and the resulting destabilizing feedback effects have been extensively studied without the inclusion of interbank liability networks; see e.g., [70, 80, 90, 94, 96, 108]. We refer to e.g., [95, 189] for a detailed review of the literature on fire sales. The Eisenberg-Noe model has been recently extended to integrate fire sales loss into the cascades of defaults in interbank networks; see e.g., [22, 62, 88, 92, 117, 132, 201]. More closely related to the content of this chapter, [106] extends the methods developed in [20, 23] to provide a resilience condition for the financial network in an integrated model of fire sales and default contagion in the case of inhomogeneous random graphs.

In Chapter 2, we study the pure default cascade process in the configuration model and provide central limit theorems for the final size of the default cascade and systemic risk. The proofs rely on a martingale limit theorem from [147] and are based on techniques developed in [152] (for the k-core problem), by transferring the contagion process to a death process represented by the balls-and-bins model. Note that Chapter 2 allows for different types of nodes and heterogeneous thresholds for directed random networks, thus extending [152] and [13] (which states a central limit theorem for bootstrap percolation in the configuration model).

Contributions and organization. To the best of our knowledge, we are the first to provide central limit theorems in an integrated model for fire sales and default contagion in random financial networks:

- Our primary contribution is to provide central limit theorems for the size of default cascade and fire sales loss in a stochastic heterogeneous financial network. This extends our previous central limit theorems in Chapter 2 for the pure default cascade process (without fire sales) in heterogeneous financial networks. We state various limit theorems for the total sold shares and the equilibrium price of illiquid assets in a stylized fire sales model. In particular, we show that the equilibrium prices of illiquid assets exhibit asymptotically Gaussian fluctuations.
- Moreover, by transferring the default cascade process to a death process problem, we provide limit theorems for a continuous (virtual) time default cascade process with fire sales. Note that although this chapter does not study the dynamic case, this virtual time (associated with the corresponding death process) allows us to study the equilibrium and the final state of contagion.

- Our numerical studies investigate the effect of heterogeneity in network structure and price impact function on the final size of the default cascade and fire sales loss. We find that financial networks with higher heterogeneity may have a smaller critical value for the shock beyond which a large fraction of institutions default, both with and without fire sales. On the other hand, for smaller shocks, the most heterogeneous network could be the least resilient. Surprisingly, the fire sales loss in two financial networks with high and low connectivities are very close to each other.

- We next use Monte Carlo methods to investigate systems with a finite number of institutions and compare them with our central limit theorem results. In particular, we show how our limit theorems can be used to construct confidence intervals for the size of contagion and fire sales loss.

- We also provide the extension of our model to a financial network with multiple types of illiquid assets and state central limit theorems in this setup.

The closed-form limit theorems that we provide in a heterogeneous financial network could also serve as a mandate for regulators to collect data on those specific network characteristics and assess systemic risk via more intensive computational methods.

Outline. The chapter is organized as follows. We introduce in Section 3.2 a general model for the network of financial counterparties and describe a mechanism for default cascade in such a network, after an exogenous macroeconomic shock. We also provide a stylized model of fire sales in a financial network with a single illiquid asset and describe how the default cascade process can be transferred to a death process problem. In Section 3.3 we give our main results on limit theorems for the final size of default cascade, the total sold shares, and the equilibrium price of the illiquid asset. In particular, we show that the equilibrium price of the illiquid asset has asymptotically Gaussian fluctuations. Numerical case studies in Section 3.4 investigate the effect of heterogeneity in the network structure and price impact function on the final size of default cascade and fire sales loss. Section 3.7 concludes. Proofs of the main theorems are given in Section 3.5. Section 3.6 provides the extension of our model to a financial network with multiple types of illiquid assets. We provide central limit theorems for default cascade with fire sales in this setup.

Notation. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If $c \in \mathbb{R}$ is a constant, we write $X_n \xrightarrow{p} c$ to denote that X_n converges in probability to c that is, for any $\epsilon > 0$, we have $\mathbb{P}(|X_n - c| > \epsilon) \to 0$ as $n \to \infty$. We write $X_n \xrightarrow{d} X$ to denote that X_n converges in distribution to X. Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers going to infinity as $n \to \infty$. We write $X_n = o_p(a_n)$, if $|X_n|/a_n \xrightarrow{p} 0$. If E_n is a measurable subset of Ω, for any $n \in \mathbb{N}$, we say that the sequence $\{E_n\}_{n\in\mathbb{N}}$ occurs with high probability (w.h.p.) or almost surely (a.s.) if $\mathbb{P}(E_n) = 1 - o(1)$, as $n \to \infty$. We denote by $\mathsf{Bin}(k, p)$ a binomial distribution corresponding to the number of successes of a sequence of k independent Bernoulli trials each having probability of success p. The notation $\mathbb{1}\{E\}$ is used for the indicator of an event E; this is 1 if E holds and 0 otherwise. We denote by $\mathcal{D}[0,\infty)$ the standard space of right-continuous functions with left limits on $[0,\infty)$ equipped with the Skorokhod topology (see e.g., [147, 156]). We shall suppress the dependence of parameters on the size of the network n when it is clear from the context. We denote by $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ the set of non-negative integers.

3.2 The Model

In this section, we describe the financial network and the default cascade model in Chapter 2, extended to account for the price impact of the liquidation of illiquid assets and fire sale effects.

3.2.1 Financial network

Consider an economy $\mathcal{E}_n$ consisting of n interlinked financial institutions (banks, companies, hedge funds, etc.) denoted by $[n] := \{1, 2, \ldots, n\}$. Interbank liabilities are represented by a matrix of nominal liabilities $(\ell_{ij})_{i,j\in[n]}$, where, for two financial institutions $i, j \in [n]$, $\ell_{ij} \geqslant 0$ denotes the cash amount that bank i owes to bank j. The total nominal liabilities of bank i is $\ell_i = \sum_{j\in[n]} \ell_{ij}$, and the total outstanding receivables sum up to $a_i = \sum_{j\in[n]} \ell_{ji}$. In addition to these interbank assets and liabilities, every institution holds claims on end-users (society, households, etc.) and vice versa. The total value of claims held by end-users on bank i (deposits) is denoted by d_i, while the total value of claims held by bank i on end-users (external assets) is denoted by e_i. Bank i holds $k_i \geqslant 0$ units of a liquid asset (cash) and $\gamma_i \in [0, \gamma_{\max}]$ units of an illiquid asset. We assume that all γ_i (for all $i \in [n]$) are bounded from above by $\gamma_{\max} > 0$. Cash has a value of one, while the illiquid asset has a positive fundamental value $p_0 > 0$.

Compared to that in Chapter 2, the nominal balance sheet of bank i is then given by:

- Assets: $e_i + k_i + \gamma_i p_0 + a_i$;
- Liabilities: $d_i + \ell_i +$ nominal net worth.

In a stress testing framework, we apply a fractional shock $\epsilon_i \in [0, 1]$ to the external assets of bank i. Table 3.1 summarizes a stylized balance sheet of bank i after the shock ϵ_i. The capital of bank i after the shock, denoted by $c_i = c_i(\epsilon_i; p_0)$, satisfies

$$c_i = k_i + \gamma_i p_0 + (1 - \epsilon_i)e_i + a_i - \ell_i - d_i, \tag{3.1}$$

which represents the capacity of bank i to absorb losses while remaining solvent.

The nominal cash balance of bank i is then $k_i + (1 - \epsilon_i)e_i + a_i - d_i - \ell_i$.

Price impact of liquidations. If bank i has a negative nominal cash balance, it faces a liquidity shortfall. In this case, bank i sells some of its shares of the illiquid asset, which negatively impacts the asset's price. We model this by considering a given inverse demand function g, which determines the equilibrium price for the illiquid asset when nx units of the asset are sold within a network of size n.

External e_i	Deposits d_i
$\epsilon_i e_i$ - loss on assets	Interbank $\ell_i = \sum_{j \in [n]} \ell_{ij}$
Interbank $a_i = \sum_{j \in [n]} \ell_{ji}$	Capital c_i
Liquid k_i	
Illiquid $\gamma_i p_0$	$\epsilon_i e_i$ - loss on capital
Assets	**Liabilities**

Table 3.1: Stylized balance sheet of bank i after shock.

We impose the following moderate technical assumptions.[1]

Assumption 3.1. *Let $p_{\min} \geqslant 0$. We assume that $g : [0, \gamma_{\max}] \to [p_{\min}, p_0]$ satisfies:*

(i) $g(0) = p_0$ *(in absence of liquidations the price is given exogenously by p_0).*

(ii) $g \in \mathcal{C}^1$ *and it is a non-increasing function of $x \in [0, \gamma_{\max}]$ (the price is non-increasing with the average excess supply x).*

(iii) $g(\gamma_{\max}) = p_{\min} \geqslant 0$.

We conclude this section by presenting examples of price impact functions that satisfy the aforementioned assumptions. These examples will be further explored in our numerical experiments in Section 3.4.

Example 3.1 (Linear Price Impact function)**.** *For $y \in [0, \gamma_{\max}]$, we set:*

$$g^{\mathrm{L}}(y) = p_0 - (p_0 - p_{\min})\big(y/\gamma_{\max}\big).$$

Example 3.2 (Quadratic Price Impact function)**.** *For $y \in [0, \gamma_{\max}]$, and $\alpha > 0$, we set*

$$g^{\mathrm{Q}}_{\alpha}(y) = p_0 - (p_0 - p_{\min})\big(y/\gamma_{\max}\big)\frac{1 - \alpha\big(y/\gamma_{\max}\big)}{1 - \alpha}.$$

Example 3.3 (Exponential Price Impact function)**.** *For $y \in [0, \gamma_{\max}]$, and $\alpha > 0$, we set*

$$g^{\mathrm{E}}_{\alpha}(y) = p_0 - (p_0 - p_{\min})\frac{1 - e^{-\alpha(y/\gamma_{\max})}}{1 - e^{-\alpha}}.$$

[1] Similar to [22, 92], we assume there is an external market for this illiquid asset that can absorb the total illiquid asset holdings of the banks at a distressed price. It is beyond the scope of this chapter to endogenize both the demand function for the illiquid asset and the financial network payments.

3.2.2 Default cascade

We now introduce a model to investigate the combined effects of insolvency contagion and fire sales. Given a shock scenario $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n) \in [0,1]^n$, when bank $i \in [n]$ has a negative cash balance and the revenue generated from selling all γ_i units of the illiquid asset fails to cover the negative cash balance, bank i defaults on its interbank liabilities.

For a given price $p \in [p_{\min}, p_0]$ of the illiquid asset, we define bank i as *p-fundamentally insolvent* if its capital, after the shock and considering the price p of the illiquid asset, is negative, i.e., $c_i(\epsilon_i; p) < 0$. We denote the set of p-fundamental defaults as:

$$\mathcal{D}_0(\boldsymbol{\epsilon}; p) := \{i \in [n] : c_i(\epsilon_i; p) < 0\}. \tag{3.2}$$

We next define the pure default cascade (excluding fire sales loss) initiated by fundamentally insolvent institutions. It is important to note that, for a given shock scenario $\boldsymbol{\epsilon}$, the price of the illiquid asset can be affected by fire sales, resulting in a price $p \leqslant p_0$. This leads to a larger set of fundamentally insolvent institutions, denoted as $\mathcal{D}_0(\boldsymbol{\epsilon}; p)$. This, in turn, triggers the default contagion process.

Let us fix the shock $\boldsymbol{\epsilon}$ and the price of the illiquid asset $p \in [p_{\min}, p_0]$. We denote the recovery rate of the liability of i to j as $R_{ij} = R_{ij}(\boldsymbol{\epsilon}; p)$, in the event that bank i defaults. The matrix of recovery rates is represented by $\mathcal{R} = (R_{ij})_{i,j \in [n]}$. Since any bank i cannot pay more than its external assets $(1-\epsilon_i)e_i$ plus what it has recovered from its debtors, the recovery rates of i must satisfy the following cash-flow consistency constraints:

$$\gamma_i p + k_i + (1-\epsilon_i)e_i + \sum_{j=1}^{n} R_{ji}\ell_{ji} \geqslant \sum_{j=1}^{n} R_{ij}\ell_{ij} + d_i.$$

Similarly as in [20, 24, 93], we assume fixed recovery rates in this chapter. The model can be expanded to include random recovery rates that satisfy the above cash-flow consistency conditions, as discussed in [24]. Note that, although not explicitly stated, the fixed recovery rates in our model could be a function of both the initial shock and the price of the illiquid asset, under the condition that the recovery rates satisfy the above cash-flow consistency constraints.

Given the shock scenario $\boldsymbol{\epsilon}$, the price of the illiquid asset p, and the matrix of recovery rates $\mathcal{R}$, a default cascade is initiated by the set of p-fundamentally insolvent institutions $\mathcal{D}_0(\boldsymbol{\epsilon}; p)$, eventually reaching the equilibrium set $\mathcal{D}^\star$. This set represents financial institutions whose capital is insufficient to absorb losses and must satisfy the following fixed-point equation:

$$\mathcal{D}^\star = \mathcal{D}^\star(\boldsymbol{\epsilon}, \mathcal{R}; p) = \Big\{ i \in [n] : c_i(\epsilon_i; p) < \sum_{j \in \mathcal{D}^\star} (1 - R_{ji})\ell_{ji} \Big\}.$$

As demonstrated in [24], the above fixed-point default cascade set may have multiple solutions. The smallest fixed-point set, which corresponds to the fewest number of defaults, can be obtained by starting from $\mathcal{D}_0 = \mathcal{D}_0(\boldsymbol{\epsilon}; p)$ and defining the following at step k:

$$\mathcal{D}_k = \mathcal{D}_k(\boldsymbol{\epsilon}, \mathcal{R}; p) = \big\{i \in [n] : c_i < \sum_{j \in \mathcal{D}_{k-1}} (1 - R_{ji})\ell_{ji}\big\}. \tag{3.3}$$

The cascade ends at the first instance when k satisfies $\mathcal{D}_k = \mathcal{D}_{k-1}$. Therefore, in a financial network of size n and for a given price p of the illiquid asset, the cascade will terminate after at most $n-1$ steps. The final set of insolvent institutions is represented by $\mathcal{D}_{n-1} = \mathcal{D}_{n-1}(\boldsymbol{\epsilon}, \mathcal{R}; p)$.

3.2.3 Node classification

As detailed below, under certain regularity assumptions, we can consolidate information regarding assets (both liquid and illiquid), liabilities, post-exogenous-shock capital, and recovery rates into a single probability threshold function (which serves as a probability mass function for the threshold random variable); see [17, 20] for a similar setup.

For a given illiquid asset price p, shock scenario $\boldsymbol{\epsilon}$, and matrix of recovery rates $\mathcal{R}$, we introduce the (random) threshold $\Theta_i(p) = \Theta_i^{(n)}(p)$ for every institution $i \in [n]$. This value represents the number of defaults that bank i can endure before becoming insolvent, assuming that the order of its counterparties' defaults is random — that is, when the order of i's debtor defaults is chosen uniformly at random from all possible permutations.

Next, we consider a classification of financial institutions into a countable set of possible classes $\mathcal{X}$, which could be finite or infinite. All observable classes for institution i are encoded in $x_i = (d_i^+, d_i^-, t_i, ...) \in \mathcal{X}$, where d_i^+ denotes the in-degree (the number of institutions i is exposed to), d_i^- signifies the out-degree (the number of institutions exposed to i), and t_i represents any other institution-specific type (e.g., credit rating, seniority class, systemic importance, etc.).

To state limit theorems, we consider a sequence of economies $\{\mathcal{E}_n\}_{n\in\mathbb{N}}$, indexed by the number of institutions. The characteristics of any institution $i \in [n]$ in the economy $\mathcal{E}_n$ are represented as $x_i^{(n)} = (d_i^{+(n)}, d_i^{-(n)}, t_i^{(n)}, ...) \in \mathcal{X}$. Without loss of generality, we assume that institutions within the same class $x \in \mathcal{X}$ have the same number of creditors (denoted by d_x^-) and debtors (denoted by d_x^+). For the sake of tractability, we make the following assumption about the probability threshold functions.

Assumption 3.2. *There exists a classification of the financial institutions into a countable set of possible classes $\mathcal{X}$ such that, for each $n \in \mathbb{N}$ and all $p \in [p_{\min}, p_0]$, institutions within the same class share the same threshold distribution function (represented as $q_x^{(n)}$ for institutions in class $x \in \mathcal{X}$). Specifically, for the economy $\mathcal{E}_n, i \in [n]$ and all $\theta \in \mathbb{N}$,*

$$\mathbb{P}(\Theta_i^{(n)}(p) = \theta) = q^{(n)}_{x_i^{(n)}}(\theta; p).$$

In particular, in the network of size n, $q_x^{(n)}(0; p)$ represents the proportion of initially insolvent institutions of type $x \in \mathcal{X}$ under the given illiquid asset price p.

We denote $\mu_x^{(n)}$ as the fraction of institutions with characteristic $x \in \mathcal{X}$ in the economy $\mathcal{E}_n$. Compared to Chapter 2, to deal with the price impact, we need some smoothness assumption on the threshold distribution $q_x^{(n)}(\theta; p)$.

Assumption 3.3. *For some probability distribution functions μ and $q(.;p)$ over the set of classes $\mathcal{X}$ (independent of n), we have $\mu_x^{(n)} \to \mu_x$ and $q_x^{(n)}(\theta;p) \to q_x(\theta;p)$ as $n \to \infty$, for all $x \in \mathcal{X}, \theta = 0, 1, \ldots, d_x^+$ and $p \in [p_{\min}, p_0]$. We also assume that the empirical threshold distributions satisfy $q_x^{(n)}(\theta;p) \in \mathcal{C}^1$ and $q_x(\theta;p) \in \mathcal{C}^1$ on $p \in [p_{\min}, p_0]$. Moreover, as $n \to \infty$, $\frac{\partial q_x^{(n)}}{\partial p}(\theta;p)$ converges uniformly to $\frac{\partial q_x}{\partial p}(\theta;p)$ as a function of p for all $x \in \mathcal{X}$ and $\theta = 0, 1, \ldots, d_x^+$.*

We provide below an example of liabilities (losses) satisfying the above assumptions.

Example 3.4 (Independent random losses)**.** *Suppose the capital of each institution (post-shock) is a constant that depends on the institution's type and the price of the illiquid asset, i.e., $c_i = c_{x_i}(p)$. Let $\{L_{x,k}\}_{k=1}^{\infty}$ be a set of independent and identically distributed (i.i.d.) continuous random variables with a common cumulative distribution function (cdf) F_x and density f_x for all $x \in \mathcal{X}$.*

We then set $q_x(0;p) = \bar{q}_x$. Further, we set

$$q_x(1;p) = (1 - \bar{q}_x)\mathbb{P}(c_x(p) \leqslant L_{x,1}) = (1 - \bar{q}_x)\big(1 - F_x(c_x(p))\big),$$

and, for $\theta = 2, \ldots, d_x^+$, we set

$$\begin{aligned} q_x(\theta;p) &= (1 - \bar{q}_x)\mathbb{P}(L_{x,1} + \cdots + L_{x,\theta-1} < c_x(p) \leqslant L_{x,1} + \cdots + L_{x,\theta}) \\ &= (1 - \bar{q}_x)\int_0^{c_x(p)} f^{\star k}(\nu)\big(1 - F_x(c_x(p) - \nu)\big)d\nu, \end{aligned}$$

where $f^{\star k}$ is the k-fold convolution of f. Since the capital $c_x(p)$ is smooth (in fact linear in p) for all $x \in \mathcal{X}$, then the threshold distribution is $\mathcal{C}^1$ in p for all $x \in \mathcal{X}$ and θ. In our numerical experiments in Section 3.4, we consider a Pareto distribution for losses, that is

$$f(x) = \alpha x_m^{\alpha} x^{-(\alpha+1)} \mathbb{1}\{x \geqslant x_m\},$$

for some scale and shape parameters $x_m, \alpha \in \mathbb{R}^+$.

In this chapter, we account for the possibility that an institution never defaults, i.e., it remains solvent even if all its counterparties default. This case is not considered in the pure default cascade process studied in Chapter 2. For $x \in \mathcal{X}$ and $p \in [p_{\min}, p_0]$, we set

$$q_x(\infty;p) := 1 - \sum_{\theta=0}^{d_x^+} q_x(\theta;p).$$

3.2.4 A simple model of fire sales

In our model, we assume that each instance of a default of an incoming link (originating from a defaulted neighbor) compels the host institution (creditor) to liquidate a portion of its asset holdings. This mechanism is introduced to reflect the institution's need to adhere to regulatory constraints, such as maintaining prescribed leverage ratios. Given the structure of the default cascade death process as outlined in Algorithm 1, we make the following assumption: each time an institution is faced with a defaulted neighbor, the number of shares it sells off are independent random variables. The distribution of these random variables depends on the host institution's type, its threshold level, and the current (equilibrium) price of the illiquid asset. We acknowledge that in reality, defaults and subsequent fire sales might not occur instantaneously. However, for the purpose of our model, we adopt a conservative approach, following the literature such as [22, 92, 133]. We assume that all institutions may only sell their assets at the final equilibrium price. This assumption is meant to capture the drastic drop in asset prices that often occurs in financial crises, thereby intensifying the feedback loop of defaults and fire sales.

Remark 3.5. *As previously discussed, institutions often need to liquidate a portion of their assets in response to losses, driven by the obligation to comply with market regulations and constraints. These sales are typically proportional to the ratio of loss to capital. A non-decreasing function $\rho : [0, \infty) \to [0, 1]$ is employed in [107] to represent this proportion of liquidation. For example, an institution $i \in [n]$ experiencing a loss L_i will liquidate $\gamma_i \rho(L_i/c_i)$ of its asset shares. Bounds are provided in [107] for the total shares sold and equilibrium price of illiquid assets. In our context, formulating central limit theorem results with a generic function ρ is complex, given the cumulative nature of fire sales. With a generic non-decreasing function ρ, each loss incurred from a defaulted neighbor could lead to different liquidation amounts, even if the received losses are equal. To simplify this issue and facilitate the study of central limit theorems for liquidation amounts and equilibrium prices after fire sales, we opt for a linear sales function in our model. Consequently, equivalent losses result in identical liquidation amounts, thus enabling us to model liquidations as independent random variable, with distribution depending on each institution's type and threshold. Another specific case studied in [107] is the sales function $\rho(u) = \mathbb{1}_{\{u \geqslant 1\}}$, which indicates complete liquidation at default; the total liquidation amount then depends only on the total number of defaults. This case can be directly linked to Chapter 2, where we study the limit theorems for default contagion without considering fire sales.*

We now provide a mathematical exposition of our fire sales model, beginning by establishing some notation summarized in Table 3.2.

For a fixed price $p \in [p_{\min}, p_0]$, we use $D^{(n)}_{x,\theta}(t;p)$ to represent the total number of defaulted institutions of type x and threshold θ at time t. Consequently, the total number of defaulted institutions of type $x \in \mathcal{X}$ at time t is given by $D^{(n)}_{x}(t;p) = \sum_\theta D^{(n)}_{x,\theta}(t;p)$. We also recall that $S^{(n)}_{x,\theta,\ell}(t;p)$ is used to denote the number of solvent institutions of type $x \in \mathcal{X}$, threshold $\theta \in \mathbb{N}$, and with ℓ defaulted neighbors at time t.

For $x \in \mathcal{X}, p \in [p_0, p_{\min}]$ and $\theta = 1, \ldots, d^+_x$, we define

$$I^{(n)}_{x,\theta}(t;p) := \theta D^{(n)}_{x,\theta}(t;p) + \sum_{\ell=1}^{\theta-1} \ell S^{(n)}_{x,\theta,\ell}(t;p).$$

This represents the total number of liquidations for institutions of type $x \in \mathcal{X}$ and threshold θ up until time t. The first term indicates that a defaulted institution of type x and threshold $\theta \geqslant 1$ will need to liquidate θ times. Additionally, a solvent institution with ℓ defaulted neighbors would need to liquidate ℓ times before time t.

We also consider institutions that never default under such a shock scenario, even if all their counterparties default. Let $S^{(n)}_{x,\infty,\ell}(t;p)$ denote the number of institutions of type x with a threshold greater than d^+_x (thus, never defaulting), and ℓ defaulted neighbors at time t. We then define

$$I^{(n)}_{x,\infty}(t;p) := \sum_{\ell=1}^{d^+_x} \ell S^{(n)}_{x,\infty,\ell}(t;p),$$

which represents the total number of liquidations from institutions of type $x \in \mathcal{X}$ that never default, up until time t.

We assume that the quantity of liquidation for each institution that is initially defaulted, of type $x \in \mathcal{X}$, is a constant value, represented by $\bar{\gamma}_x$. The symbol $D^{(n)}_{x,0}(0;p) = nq^{(n)}_x(0;p)$ is used to represent the number of initially insolvent institutions belonging to type $x \in \mathcal{X}$.

For every type $x \in \mathcal{X}$ and threshold value θ within the set $\{1,\ldots,d^+_x\} \cup \{\infty\}$, we define a series $\{L^{(i)}_{x,\theta}(p)\}_{i=1}^{\infty}$. This represents a sequence of independent and identically distributed (i.i.d.) positive bounded random variables that follow the same distribution, denoted as $F_{x,\theta}(.;p)$. Specifically, $L^{(i)}_{x,\theta}(p)$ represents the quantity of illiquid asset sold upon the i-th default occurring to institutions of type x with threshold θ. Similarly, $L^{(i)}_{x,\infty}(p)$ denotes the quantity of illiquid asset sold upon the i-th default occurring to institutions of type x that never default (those with threshold larger than d^+_x). These random variables, given the price $p \in [p_{\min}, p_0]$ of the illiquid asset, have a mean value symbolized by $\bar{\ell}_{x,\theta}(p)$ and a variance represented by $\varsigma^2_{x,\theta}(p)$, which satisfy the following assumption.

Assumption 3.4. *The mean $\bar{\ell}_{x,\theta}(p)$ and the variance $\varsigma^2_{x,\theta}(p)$ of the shares sold for each liquidation are continuous functions of p, and this holds for every type $x \in \mathcal{X}$ and threshold $\theta \in \{0,1,\ldots,d^+_x\} \cup \{\infty\}$.*

The total number of illiquid asset shares sold by time t (for a given price p of illiquid asset) can be expressed as

$$\Gamma_n(t;p) := \sum_{x\in\mathcal{X}} \Big(\bar{\gamma}_x D^{(n)}_{x,0}(p) + \sum_{\theta=1}^{d^+_x} Y^{(n)}_{x,\theta}(t;p), + Y^{(n)}_{x,\infty}(t;p)\Big), \tag{3.4}$$

where

$$Y^{(n)}_{x,\theta}(t;p) := \sum_{i=1}^{I^{(n)}_{x,\theta}(t;p)} L^{(i)}_{x,\theta}(p) \quad \text{and} \quad Y^{(n)}_{x,\infty}(t;p) := \sum_{i=1}^{I^{(n)}_{x,\infty}(t;p)} L^{(i)}_{x,\infty}(p) \tag{3.5}$$

represent the total number of asset shares sold by institutions of type x with threshold θ and institutions of type x that never default, respectively, up to time t.

Finally, the total shares of the illiquid asset that have been sold under price p at the stopping time $\tau^\star_n(p)$ will be denoted as $\Gamma_n(\tau^\star_n(p);p)$.

$D_{x,0}^{(n)}(0;p)$	number of initially defaulted institutions with type $x \in \mathcal{X}$
$D_{x,\theta}^{(n)}(t;p)$	number of defaulted institutions of type x with threshold θ at time t
$S_{x,\theta,\ell}^{(n)}(t;p)$	number of solvent institutions of type x, threshold $\theta \in \{1,\ldots,d_x^+\} \cup \{\infty\}$ and with ℓ defaulted neighbors at time t
$I_{x,\theta}^{(n)}(t;p)$	total number of liquidations for institutions with type x and threshold $\theta \in \{1,\ldots,d_x^+\} \cup \{\infty\}$ up to time t
$\{L_{x,\theta}^{(i)}(p)\}_{i=1}^{\infty}$	a set of i.i.d. positive bounded random variables representing units of illiquid asset sold at each incoming default leading to institutions with type x and threshold $\theta \in \{1,\ldots,d_x^+\} \cup \{\infty\}$
$\bar{\ell}_{x,\theta}(p)$	the mean value of $L_{x,\theta}^{(i)}(p)$
$\varsigma_{x,\theta}^2(p)$	the variance of $L_{x,\theta}^{(i)}(p)$
$Y_{x,\theta}^{(n)}(t;p)$	total shares of illiquid asset sold up to time t for institutions with type $x \in \mathcal{X}$ and threshold $\theta \in \{1,\ldots,d_x^+\} \cup \{\infty\}$
$\bar{\gamma}_x$	the constant value of liquidation for each initially defaulted institution with type x
$\Gamma_n(t;p)$	total shares of illiquid asset sold by time t
$\kappa_n(p)$	the price of the asset given by the inverse demand function g
$p_n^\star$	the equilibrium price of the illiquid asset

Table 3.2: Overview of the fire sales model notation, under price p of the illiquid asset

Given that default contagion and fire sales occur instantly in our model, we adopt a conservative strategy. We make the assumption that the illiquid asset can only be sold at the final equilibrium price.

We establish this price via the inverse demand function, g, defined as follows:

$$\kappa_n(p) := g(\Gamma_n(\tau_n^\star(p);p)/n).$$

Since $\tau_n^\star(p)$ is not in general continuous (this is shown in the next section), the fixed point equation $p = \kappa_n(p)$ may not have a solution. This motivates us to define the *equilibrium price* of the illiquid asset as

$$p_n^\star = \sup\big\{p \in [p_{\min}, p_0] : p \leqslant \kappa_n(p)\big\}, \tag{3.6}$$

that is the highest price, lower than or equal to the one given by the inverse demand function $\kappa_n(p)$, within the range $[p_{\min}, p_0]$. This equilibrium price provides an optimal price for the illiquid asset considering the constraints of the model.

3.3 Limit Theorems

In this section, we establish limit theorems for the total sold shares and equilibrium price of the illiquid asset in the random financial network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$, which is defined in the same way as in

Chapter 2. Some assumptions are the same as those in Chapter 2, and we also incorporate results from Chapter 2 with minor adaptations to the integrated default contagion and fire sales model. For the sake of completeness, we restate them in this chapter.

3.3.1 Asymptotic magnitude of default cascade with fire sales

We assume the following regularity condition on the average degrees.

Assumption 3.5a. *We assume that as $n \to \infty$ the average degrees converges to a finite limit:*

$$\lambda^{(n)} := \sum_{x\in\mathcal{X}} d_x^+ \mu_x^{(n)} = \sum_{x\in\mathcal{X}} d_x^- \mu_x^{(n)} \longrightarrow \lambda := \sum_{x\in\mathcal{X}} d_x^+ \mu_x \in (0,\infty).$$

For $z \in [0,1]$ and $p \in [p_0, p_{\min}]$, we define the functions:

$$f_S(z;p) := \sum_{x\in\mathcal{X}} \mu_x \Big[\sum_{\theta=1}^{d_x^+} q_x(\theta;p)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big) + q_x(\infty;p)\Big], \quad f_D(z;p) = 1 - f_S(z;p), \tag{3.7}$$

$$f_W(z;p) := \lambda z - \sum_{x\in\mathcal{X}} \mu_x d_x^- \Big[\sum_{\theta=1}^{d_x^+} q_x(\theta;p)\beta\big(d_x^+, z, d_x^+ - \theta + 1\big) + q_x(\infty;p)\Big]. \tag{3.8}$$

The following lemma provides the law of large numbers for the number of solvent/defaulted institutions and the total number of existing white outgoing half-edges (controlling the contagion stopping time) at any time t in the economy $\mathcal{E}_n$ satisfying the above regularity assumptions. The lemma extends Theorem 2.1 in Chapter 2 by allowing that some institutions never default (the institutions with ∞ threshold). This theorem is proved in Chapter 2 for fixed threshold distribution and can be applied for a fix $p \in [p_{\min}, p_0]$. Note that here the limiting functions f_W and f_S are slightly different from those in Chapter 2; see Section 3.5.5 for discussion.

Lemma 3.6. *Let $\tau_n \leqslant \tau_n^\star(p)$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. For all $x \in \mathcal{X}, \theta = 1, \dots, d_x^+$ and $\ell = 0, \dots, \theta - 1$, we have (as $n \to \infty$)*

$$\sup_{t\leqslant\tau_n} \Big| \frac{S^{(n)}_{x,\theta,\ell}(t;p)}{n} - \mu_x q_x(\theta;p) b\left(d_x^+, 1-e^{-t}, \ell\right) \Big| \xrightarrow{p} 0.$$

Moreover, as $n \to \infty$,

$$\sup_{t\leqslant\tau_n} \Big|\frac{S_n(t;p)}{n} - f_S(e^{-t};p)\Big| \xrightarrow{p} 0, \quad \sup_{t\leqslant\tau_n} \Big|\frac{D_n(t;p)}{n} - f_D(e^{-t};p)\Big| \xrightarrow{p} 0,$$

and the number of white outgoing defaulted half-edges satisfies

$$\sup_{t\leqslant\tau_n} \Big|\frac{W_n(t;p)}{n} - f_W(e^{-t};p)\Big| \xrightarrow{p} 0. \tag{3.9}$$

The lemma establishes that for a given price p of the illiquid asset, as the size of the network n approaches infinity, key quantities that describe the state of the default cascade in the financial network converge in probability to their expected values. This is particularly important as it outlines the relationship between the final proportion of defaults and the structural composition of the financial network after shocks. These structural elements include the distribution of types, thresholds, and degrees, in conjunction with the price of the illiquid asset. Such insights are pivotal for evaluating systemic risk within large, complex financial networks, as well as for assessing their susceptibility to cascading defaults.

We subsequently use Lemma 3.6 to provide a limit theorem for the cumulative sold shares at price $p \in [p_{\min}, p_0]$ up until time t. To this end, we define the following functions that we demonstrate serve as the limiting functions of $I^{(n)}_{x,\theta}(e^{-t}; p)/n$, $I^{(n)}_{x,\infty}(e^{-t}; p)/n$, and $\Gamma_n(e^{-t}; p)/n$, respectively:

$$f_{x,\theta}(z;p) := \mu_x q_x(\theta;p)\big(\theta - \sum_{\ell=d_x^+-\theta+1}^{d_x^+} \beta(d_x^+, z, \ell)\big), \qquad f_{x,\infty}(z;p) := (1-z)\mu_x q_x(\infty;p) d_x^+, \tag{3.10}$$

and,

$$f_\Gamma(z;p) := \sum_{x\in\mathcal{X}} \Big(\mu_x \bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p) f_{x,\theta}(z;p) + \bar{\ell}_{x,\infty}(p) f_{x,\infty}(z;p)\Big). \tag{3.11}$$

We have the following law of large numbers for the aggregate volume of sold shares at a particular time t, given the price $p \in [p_{\min}, p_0]$.

Theorem 3.7. *Let $\tau_n \leqslant \tau_n^\star$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Then, as $n \to \infty$ and for all $p \in [p_{\min}, p_0]$,*

$$\sup_{t \leqslant \tau_n} \Big| \frac{\Gamma_n(t;p)}{n} - f_\Gamma(e^{-t};p) \Big| \xrightarrow{p} 0.$$

Proof. See Appendix 3.5.2. □

Recall that the stopping time $\tau_n^\star(p)$ is defined as the first time when $W_n(\tau_n^\star; p)$ becomes negative. We introduce $z^\star(p)$, which is the supremum of z values in the range $[0,1]$ for which $f_W(z;p)$ (as defined in (3.8)) equals zero:

$$z^\star(p) := \sup\{z \in [0,1] : f_W(z;p) = 0\}. \tag{3.12}$$

Given that $f_W(1;p) \geqslant 0$ and $f_W(0;p) \leqslant 0$ for any p in $[p_{\min}, p_0]$, and that $f_W(z;p)$ is a continuous function, $z^\star(p)$ is well-defined. We have the following lemma from Chapter 2, which discusses the asymptotic behavior of $\tau_n^\star(p)$.

Lemma 3.8. *For any fixed $p \in [p_0, p_{\min}]$, we have (as $n \to \infty$):*

(i) If $z^\star(p) = 0$ then $\tau_n^\star(p) \xrightarrow{p} \infty$.

(ii) If $z^\star(p) \in (0,1]$ and $z^\star(p)$ is a stable solution, i.e., $f'_W(z^\star;p) > 0$, then $\tau^\star_n(p) \xrightarrow{p} -\ln z^\star(p)$.

Applying Theorem 3.7 and Lemma 3.8, we establish the following limit theorem for the final sold shares of illiquid assets $\Gamma_n(\tau^\star_n;p)$.

Theorem 3.9. *For any fixed $p \in [p_{\min}, p_0]$, the final number of sold shares satisfies:*

(i) If $z^\star(p) = 0$ then asymptotically almost all institutions default after shock and (as $n \to \infty$)

$$\frac{\Gamma_n(\tau^\star_n;p)}{n} \xrightarrow{p} \sum_{x\in\mathcal{X}} \mu_x\Big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta;p)\Big).$$

(ii) If $z^\star(p) \in (0,1]$ and $z^\star(p)$ is a stable solution, i.e., $f'_W(z^\star(p);p) > 0$, then, as $n \to \infty$,

$$\frac{\Gamma_n(\tau^\star_n;p)}{n} \xrightarrow{p} f_\Gamma(z^\star(p);p).$$

Proof. See Appendix 3.5.3. □

The theorem essentially establishes a relationship between the final sold shares of illiquid assets and the key characteristics of the financial network. It achieves this by detailing how the final sold shares converge (as the network size grows to infinity) to a limiting value that is a function of the network's structure and the average amount of liquidation, based on the type and threshold of each institution.

Since g is continuous according to Assumption 3.1, we can employ the continuous mapping theorem to determine the convergence of $\kappa_n(p)$. This comes as a direct corollary of Theorem 3.7, thus providing insights into the asymptotic behavior of the price $\kappa_n(p)$, defined as $g(\Gamma_n(\tau^\star_n(p);p)/n)$.

Corollary 3.10. *For any fixed $p \in [p_{\min}, p_0]$ and as $n \to \infty$, the price $\kappa_n(p)$, determined by the inverse demand function, satisfies:*

(i) If $z^\star(p) = 0$ then asymptotically almost all institutions default after shock. Consequently,

$$\kappa_n(p) \xrightarrow{p} g\Big(\sum_{x\in\mathcal{X}} \mu_x\big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta;p)\big)\Big).$$

(ii) If $z^\star(p) \in (0,1]$ and $z^\star(p)$ is a stable solution, i.e., $f'_W(z^\star(p);p) > 0$, then

$$\kappa_n(p) \xrightarrow{p} g\big(f_\Gamma(z^\star(p);p)\big).$$

We proceed by presenting a limit theorem for the equilibrium price after a shock, as defined by Equation (3.6). Corollary 3.10 serves as the motivation for introducing the following notation:

$$\bar{p} := \sup\big\{p \in [p_{\min}, p_0] : p \leqslant g\big(f_\Gamma(z^\star(p);p)\big)\big\}. \tag{3.13}$$

We say that $\bar{p}$ is a *stable* fixed point solution if it satisfies either $\bar{p} = p_{\min}$ or, in the case $\bar{p} \in (p_{\min}, p_0]$, there exists an $\epsilon > 0$ such that $p < g\big(f_\Gamma(z^\star(p);p)\big)$ for all $p \in (\bar{p} - \epsilon, \bar{p})$.

Theorem 3.11. *As $n \to \infty$, the equilibrium price $p_n^\star$ (as defined in (3.6)) satisfies:*

(i) If $z^\star(\bar{p}) = 0$ and $\bar{p}$ is a stable solution, then the equilibrium price $p_n^\star$ converges to $\bar{p}$ in probability. In this case, $\bar{p}$ is the largest solution of the fixed point equation

$$p = g\Big(\sum_{x\in\mathcal{X}} \mu_x\big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta;p)\big)\Big).$$

(ii) If $z^\star(\bar{p}) \in (0,1]$ is a stable solution of $f_W(z;\bar{p})$, i.e., $\frac{\partial f_W}{\partial z}(z^\star;\bar{p}) > 0$, and $\bar{p}$ is a stable solution of Equation (3.13), then as $n \to \infty$, we have

$$p_n^\star \xrightarrow{p} \bar{p}.$$

Proof. See Appendix 3.5.4. □

The theorem establishes a link between the equilibrium price of a liquid asset and key characteristics of a financial network, including the post-shock type and threshold distribution and the average amount of liquidation based on each institution's type and threshold. This connection provides valuable insights for regulators to evaluate the propagation of shocks and their impact on the overall stability of the financial system. By understanding these limit theorems, regulators can identify vulnerabilities within the network and implement measures to mitigate systemic risk effectively.

Remark 3.12. *The limit theorems presented have practical implications for establishing a resilience condition for default cascades in random financial networks. Specifically, using the notation introduced earlier, a financial network is considered resilient if, starting from a small fraction ϵ of institutions representing fundamental defaults, the limit of $z^\star(\bar{p})$ approaches zero as ϵ approaches zero. This resilience condition indicates that the network has the ability to withstand small shocks, as the impact of initial defaults does not result in widespread propagation throughout the financial network. We refer to [12, 20, 24, 106] for discussions on the resilience conditions for default cascades in random financial networks.*

3.3.2 Asymptotic normality of default cascade with fire sales

In order to investigate the central limit theorems, we need to restrict our attention to a category characterized by 'more sparse' or 'diluted' networks. Specifically, we consider the random financial network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ and make the assumption that the degree sequences satisfy the following moment condition.[2]

Assumption 3.5b. *We assume that for every constant $A > 1$, we have*

$$\sum_{i=1}^{n} A^{d_i^+} = n\sum_{x\in\mathcal{X}} \mu_x^{(n)} A^{d_x^+} = O(n) \quad \textit{and} \quad \sum_{i=1}^{n} A^{d_i^-} = n\sum_{x\in\mathcal{X}} \mu_x^{(n)} A^{d_x^-} = O(n).$$

[2]The finite moments condition is commonly assumed to establish the central limit theorems for diffusion processes in random graphs, see e.g., [152] for the k-core and [?] for bootstrap percolation.

In relation to the limit functions introduced in (3.7) and (3.8), for $z \in [0,1]$ and $p \in [p_0, p_{\min}]$, we define the functions $f_S^{(n)}(z;p), f_D^{(n)}(z;p)$ and $f_W^{(n)}(z;p)$ as

$$f_S^{(n)}(z;p) := \sum_{x\in\mathcal{X}} \mu_x^{(n)} \Big[\sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta;p) \beta\big(d_x^+, z, d_x^+ - \theta + 1\big) + q_x^{(n)}(\infty;p) \Big], \quad f_D^{(n)}(z;p) = 1 - f_S^{(n)}(z;p),$$

$$f_W^{(n)}(z;p) := \lambda^{(n)} z - \sum_{x\in\mathcal{X}} \mu_x^{(n)} d_x^- \Big[\sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta;p) \beta\big(d_x^+, z, d_x^+ - \theta + 1\big) + q_x^{(n)}(\infty;p) \Big].$$

For convenience and to facilitate formulation, we introduce a time transformation of the functions $\widehat{f}_i^{(n)}(t;p)$ by the relation

$$\widehat{f}_i^{(n)}(t;p) = f_i^{(n)}(e^{-t};p), \text{ for } i \in \{S, D, W\}.$$

The following lemma, drawn from Chapter 2, provides the central limit theorem for the number of solvent institutions, the number of defaulted institutions, and the total number of existing white outgoing half-edges, which control the contagion stopping time. This is valid at any time t in the economy $\mathcal{E}_n$ which satisfies the aforementioned regularity assumptions.

Lemma 3.13. *Let $\tau_n \leqslant \tau_n^\star(p)$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$.*

(i) For all $x \in \mathcal{X}, \theta = 1, \ldots, d_x^+, \ell = 0, \ldots, \theta - 1$ and jointly in $\mathcal{D}[0,\infty)$,

$$n^{-1/2} \left(S_{x,\theta,\ell}^{(n)}(t \wedge \tau_n; p) - n\mu_x^{(n)} q_x^{(n)}(\theta;p) b\big(d_x^+, 1 - e^{-(t\wedge\tau_n)}, \ell\big) \right) \xrightarrow{d} \mathcal{Z}_{x,\theta,\ell}(t \wedge t_0; p),$$

where $\mathcal{Z}_{x,\theta,\ell}(t;p)$ is a Gaussian process with mean 0 and variance $\sigma_{x,\theta,\ell}(t;p)$[3].

(ii) For all $i \in \{S, D, W\}$, as $n \to \infty$ and jointly in $\mathcal{D}[0,\infty)$,

$$n^{-1/2} \left(i_n(t \wedge \tau_n; p) - n\widehat{f}_i^{(n)}(t \wedge \tau_n; p) \right) \xrightarrow{d} \mathcal{Z}_i(t \wedge t_0; p), \tag{3.14}$$

where $\{Z_i\}$ are continuous Gaussian processes on $[0, t_0]$ with mean 0. The variance of $\mathcal{Z}_W$ denoted by $\sigma_W(e^{-t};p) := \mathrm{Var}(\mathcal{Z}_W(t;p))$, is given by (3.21).

The lemma demonstrates that the final size of a default cascade exhibits asymptotically Gaussian fluctuations. Furthermore, it relates the variance of these fluctuations to the characteristics of the financial network. In particular, it ties the variance to the post-shock type and threshold distribution.

In relation to (3.10) and (3.11), we now define the following functions which can be interpreted as the limiting functions of $I_{x,\theta}^{(n)}(e^{-t};p)/n$, $I_{x,\infty}^{(n)}(e^{-t};p)/n$ and $\Gamma_n(e^{-t};p)/n$, respectively:

$$f_{x,\theta}^{(n)}(z;p) := \mu_x^{(n)} q_x^{(n)}(\theta;p) \big(\theta - \sum_{\ell=d_x^+-\theta+1}^{d_x^+} \beta(d_x^+, z, \ell)\big), \quad f_{x,\infty}^{(n)}(z;p) := (1-z)\mu_x^{(n)} q_x^{(n)}(\infty;p) d_x^+,$$

[3] The explicit form of $\sigma_{x,\theta,\ell}(t;p)$ is provided in Chapter 2.

and,

$$f_\Gamma^{(n)}(z;p) := \sum_{x\in\mathcal{X}} \Big(\mu_x^{(n)} \bar{\gamma}_x q_x^{(n)}(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p) f_{x,\theta}^{(n)}(z;p) + \bar{\ell}_{x,\infty}(p) f_{x,\infty}^{(n)}(z;p) \Big).$$

The time-transformed versions of the above functions are defined as follows:

$$\widehat{f}_{x,\theta}^{(n)}(t;p) := f_{x,\theta}^{(n)}(e^{-t};p), \qquad \widehat{f}_\Gamma^{(n)}(t;p) := f_\Gamma^{(n)}(e^{-t};p),$$

(This transformation applies equally to any other relevant functions).

By using Lemma 3.13, we prove the following central limit theorem for the total sold shares.

Theorem 3.14. *Let $\tau_n \leqslant \tau_n^\star$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Then for any fixed $p \in [p_{\min}, p_0]$ and $t > 0$, as $n \to \infty$,*

$$n^{-1/2}(\Gamma_n(t \wedge \tau_n;p) - n\widehat{f}_\Gamma^{(n)}(t \wedge \tau_n;p)) \xrightarrow{d} \mathcal{Z}_\Gamma(t \wedge t_0;p), \tag{3.15}$$

where $\mathcal{Z}_\Gamma(t;p)$ is a Gaussian random variable with mean 0 and variance

$$\Psi(t;p) := \operatorname{Var}(\mathcal{Z}_\Gamma(t;p)),$$

where the form of $\Psi(t;p)$ is given by (3.25).

Proof. See Appendix 3.5.6. □

In order to state the central limit theorem for the final total of sold shares, we will use the following notation for $i \in \{\Gamma, W\}$:

$$f_i^1(z;p) := \frac{\partial f_i}{\partial z}(z;p), \qquad f_i^2(z;p) := \frac{\partial f_i}{\partial p}(z;p),$$

and,

$$f_i^{1,(n)}(z;p) := \frac{\partial f_i^{(n)}}{\partial z}(z;p), \qquad f_i^{2,(n)}(z;p) := \frac{\partial f_i^{(n)}}{\partial p}(z;p).$$

Remark 3.15. *Under Assumption 3.3, the bivariate functions $f_W(z;p)$, $f_W^{(n)}(z;p)$, $f_\Gamma(z;p)$, and $f_\Gamma^{(n)}(z;p)$ all possess continuous first order partial derivatives with respect to both z and p. Moreover, for any pair $(z,p) \in [0,1] \times [p\text{min}, p_0]$, as $n \to \infty$, we have:*

$$f_i^{1,(n)}(z;p) \to f_i^1(z;p) \quad \text{and} \quad f_i^{2,(n)}(z;p) \to f_i^2(z;p).$$

For a fixed z, the convergence with respect to p is uniform for all $f_i^{(n)}$ and their p derivatives. As indicated in Chapter 2, under Assumption 3.5b, these convergences extend to uniformity with respect to z, along with all derivatives with respect to z, for any fixed price p. Therefore, under Assumption 3.5b, the convergences are uniform with respect to both variables z and p.

In connection with (3.12), for the network of size n, we define

$$z_n^\star(p) := \sup\{z \in [0,1] : f_W^{(n)}(z;p) = 0\}. \tag{3.16}$$

Subsequently, we define $t^\star(p) := -\ln z^\star(p)$ and $t_n^\star(p) := -\ln z_n^\star(p)$.

Building upon Lemma 3.13 and Theorem 3.14, we present the following theorem concerning the asymptotic normality of the final total of sold shares.

Theorem 3.16. *For a fixed $p \in [p_{\min}, p_0]$, if $z^\star(p) \in (0,1)$ is a stable solution, i.e., $\alpha(p) := f_W^1(z^\star(p);p) > 0$, then as $n \to \infty$, the final total of sold shares satisfy:*

$$n^{-1/2}(\Gamma_n(\tau_n^\star;p) - n\widehat{f}_\Gamma^{(n)}(t_n^\star(p);p)) \xrightarrow{d} \mathcal{Z}_\Gamma(t^\star(p);p) - \alpha(p)^{-1} f_\Gamma^1(z^\star(p);p)\mathcal{Z}_W(t^\star(p);p),$$

where $\mathcal{Z}_\Gamma$ and $\mathcal{Z}_W$ represent Gaussian random variables with a mean of 0, as outlined in Theorem 3.14 and Lemma 3.13, respectively.

Proof. See Appendix 3.5.7. □

The theorem demonstrates that the final total of sold shares exhibits asymptotically Gaussian fluctuations. Furthermore, it associates the variance of these fluctuations with the characteristics of the financial network. Specifically, it links the variance to both the post-shock type and threshold distribution, and to the variability in the liquidation amounts, each of which is contingent on the specific type and threshold of each institution.

Remark 3.17. *Note that Theorem 3.16 cannot be applied in the boundary cases $z^\star(p) = 1$ or $z^\star(p) = 0$. When $z^\star(p) = 1$, the initial shock will not trigger a default cascade in the network. The variance of the asymptotic Gaussian in this situation arises solely from the randomness of the initial defaults, not from any randomness introduced by the default cascade. On the other hand, if $z^\star(p) = 0$, as per Theorem 3.9, almost all institutions default after a shock in the asymptotic limit, i.e., as $n \to \infty$,*

$$\frac{\Gamma_n(\tau_n^\star;p)}{n} \xrightarrow{p} \sum_{x\in\mathcal{X}} \mu_x\Big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta;p)\Big).$$

In this situation, the $z_n^\star(p)$ values are always to the right of 0 for all n, meaning they cannot be negative. Investigating a critical window for this case would be substantially more complex, and we leave this for future work.

As a corollary of Theorem 3.16, we can derive the following theorem regarding the price as determined by the inverse demand function $\kappa_n(p) := g(\Gamma_n(\tau_n^\star(p);p)/n)$.

Theorem 3.18. *For any fixed $p \in [p_{\min}, p_0]$, if $z^\star(p) \in (0,1)$ and $z^\star(p)$ is a stable solution, i.e., $\alpha(p) := f_W^1(z^\star(p);p) > 0$, then as $n \to \infty$, the price $\kappa_n(p)$ as determined by the inverse demand function satisfies*

$$n^{1/2}\big(\kappa_n(p) - g\big(\widehat{f}_\Gamma^{(n)}(t_n^\star(p);p)\big)\big) \xrightarrow{d} g'\big(f_\Gamma(z^\star(p);p)\big)\Big[\mathcal{Z}_\Gamma(t^\star(p);p) - \alpha(p)^{-1} f_\Gamma^1(z^\star(p);p)\mathcal{Z}_W(t^\star(p);p)\Big],$$

where g' denotes the first derivative of g.

Proof. See Appendix 3.5.8. □

We now present a central limit theorem for the equilibrium price post-shock, as defined by Equation (3.6). In connection to (3.13), for a network of size n, we define:

$$\bar{p}_n := \sup\big\{p \in [p_{\min}, p_0] : p \leqslant g\big(f_\Gamma^{(n)}(z_n^\star(p); p)\big)\big\}. \tag{3.17}$$

Recall that $\bar{p}$ is a *stable* fixed point solution if it satisfies either $\bar{p} = p_{\min}$ or, in the case $\bar{p} \in (p_{\min}, p_0]$, there exists an $\epsilon > 0$ such that $p < g\big(f_\Gamma(z^\star(p); p)\big)$ for all $p \in (\bar{p} - \epsilon, \bar{p})$.

Theorem 3.19. *If $z^\star(\bar{p}) \in (0, 1)$ is a stable solution of $f_W(z; \bar{p}) = 0$, i.e., $\alpha(\bar{p}) := f_W^1(z^\star; \bar{p}) > 0$, and $\bar{p}$ is a stable solution of (3.13), then as $n \to \infty$, the equilibrium price satisfies*

$$n^{1/2}(p_n^\star - \bar{p}_n) \xrightarrow{d} -\rho^{-1}(\bar{p})\mathcal{Z}_V(\bar{p}),$$

where

$$\rho(p) := 1 - g'\big(f_\Gamma(z^\star(p); p)\big)\Big[-f_\Gamma^1(z^\star(p); p)\alpha(p)^{-1} f_W^2(z^\star(p); p) + f_\Gamma^2(z^\star(p); p)\Big],$$

and,

$$\mathcal{Z}_V(p) := -g'\big(f_\Gamma(z^\star; p)\big)\Big[\mathcal{Z}_\Gamma(t^\star(p); p) - \alpha(p)^{-1} f_\Gamma^1(z^\star; p)\mathcal{Z}_W(t^\star(p); p)\Big]$$

is a Gaussian random variable with mean 0.

Proof. See Appendix 3.5.9. □

The theorem establishes that the final equilibrium price exhibits asymptotically Gaussian fluctuations. Additionally, it establishes a connection between the variance of these fluctuations and the characteristics of the financial network and the inverse demand function. Specifically, the variance is linked to the post-shock type and threshold distribution, as well as the variability in the liquidation amounts, both of which depend on the specific type and threshold of each institution. To highlight the practical significance of central limit theorems, we proceed to numerically investigate the asymptotic variance for fire sales loss in the following section.

3.4 Numerical Experiments

Empirical studies on the network topology of banking systems reveal a wide range of structures, including centralized networks as shown in [176], core-periphery structures explored in [99, 122, 166] as well as scale-free structures discussed in [69, 93]. In this section, we examine the impact of heterogeneity in network structure and price impact function on the final size of default cascades and fire sale losses.

3.4.1 Numerical set-up

In our numerical experiments, we make the assumption that the in-degree and out-degree of each institution are equal, denoted as $d_x^+ = d_x^- = d_x$ for all $x \in \mathcal{X}$. We normalize the price of the illiquid asset to fall between $p_{\min} = 1$ and $p_0 = 2$. Additionally, we assume that institutions of the same type or class share the same capital structure. To describe the capital structure of institutions with type $x \in \mathcal{X}$, we employ the capital vector $\mathbf{h}_x$, given by (see Table 3.1)

$$\mathbf{h}_x := [\gamma_x \quad k_x + a_x \quad \ell_x + d_x \quad e_x].$$

In our stress testing framework, we make the assumption that the initial fraction of defaults is fixed across all classes, denoted as $q_x(0; p) = \epsilon$ for all $x \in \mathcal{X}$. For the sake of illustration, we set $\epsilon_i = \epsilon$ uniformly across all institutions, resulting in each initially solvent institution experiencing a loss of a fraction $\epsilon \in [0, 1]$ of its external assets. When an institution defaults, its creditors face losses, which are assumed to be i.i.d. random variables following a Pareto distribution. The scale and shape parameters of the Pareto distribution are type-dependent and denoted as x_m and $\alpha \in \mathbb{R}^+$, respectively, to be specified. The threshold distributions can then be calculated as outlined in Example 3.4.

We consider a scenario where initially defaulted institutions liquidate all of their shares of illiquid assets. As a result, the mean liquidation fraction for institutions with the same type becomes equal to their capital allocation parameter, i.e., $\bar{\gamma}_x = \gamma_x$. The mean liquidation amounts follow a linear relationship given by

$$\bar{\ell}_{x,\theta}(p) = \frac{\gamma_x}{p\theta} \quad \text{for} \quad \theta = 1, \dots, d_x,$$

and we set $\bar{\ell}_{x,\infty}(p) = \frac{\gamma_x}{2pd_x}$ for all $p \in [1, 2]$. This specification allows us to determine the mean liquidation amounts based on the type of institution and the price of the illiquid asset.

We shall consider three different price impact functions, each with specific forms as provided in Examples 3.1, 3.2, and 3.3. The functions are defined as follows:

- Linear price impact (LPI): $g^L(y) = 2 - (y/\gamma_{\max})$;
- Quadratic price impact (QPI): $g^Q_\infty(y) = 2 - (y/\gamma_{\max})^2$;
- Exponential price impact (EPI): $g^E_1(y) = 2 - \frac{1-e^{-(y/\gamma_{\max})}}{1-e^{-1}}$.

These functions are defined for $y \in [0, \gamma_{\max}]$. It is important to note that the LPI function decreases at a constant rate for all values of y. In contrast, the QPI function initially drops slowly (for small y) and then decreases faster as y increases. Conversely, the EPI function decreases rapidly at the beginning and then gradually slows down as y increases.

To quantify the extent of losses incurred by the financial system due to fire sales triggered by the exogenous shock ϵ, we utilize the Fire Sales Losses (FSL) indicator, defined as

$$\text{FSL}(\epsilon) = \frac{p_0 - p_n^\star(\epsilon)}{p_0},$$

where $p_n^\star(\epsilon)$ represents the equilibrium price of the illiquid asset following the shock ϵ. The FSL indicator measures the proportion of loss in the equilibrium price relative to the initial price p_0. A higher value of FSL indicates a greater impact of fire sales on the financial system, resulting in larger losses.

3.4.2 Regular financial networks

We consider a regular network, where all institutions are of the same type and have the same degree d, and investigate the impact of network connectivity and fire sales on the size of default cascades and the resulting losses in the financial system during a crisis. Specifically, we compare two scenarios with high and low network connectivity. One prominent finding in the financial network literature is that, for regular homogeneous financial networks, when shocks are small, higher connectivity leads to a lower risk of contagion. This has been demonstrated, for example, in [1] through a comparison of ring and complete network structures. Our findings align with this notion, as we observe that the risk of contagion in the two financial networks with high and low connectivity is very similar. However, we also examine the impact of fire sales losses in these two networks and find that the resulting losses are nearly identical, regardless of the network connectivity. This highlights the significance of considering not only the risk of contagion but also the potential losses associated with fire sales in assessing the overall stability and resilience of the financial system.

For a d-regular financial network, the limiting function of the white outgoing defaulted half-edges process can be simplified as follows:

$$f_W(z;p) = d\big(z - \sum_{\theta=1}^{d} q(\theta;p)\beta(d,z,d-\theta+1) - q(\infty;p)\big).$$

Thus, in this case, the expression for $z^\star(p)$ is given by:

$$z^\star(p) := \sup\big\{z \in [0,1] : z = \sum_{\theta=1}^{d} q(\theta;p)\beta(d,z,d-\theta+1) + q(\infty;p)\big\}.$$

Similarly, the limiting function of the total liquidation process can be simplified as

$$f_\Gamma(z;p) = \gamma q(0;p) + \frac{\gamma}{p}\sum_{\theta=1}^{d}\frac{q(\theta;p)}{\theta}\big(\theta - \sum_{\ell=d-\theta+1}^{d}\beta(d,z,\ell)\big) + \frac{\gamma}{2p}(1-z)q(\infty;p).$$

Recall that from Theorem 3.11, the equilibrium price of the illiquid asset after shock ϵ is denoted as $\bar{p} = \bar{p}(\epsilon)$ and is given by (3.13). The limiting fire sales loss can then be expressed as

$$\mathrm{FSL}(\epsilon) = \frac{p_0 - \bar{p}(\epsilon)}{p_0}.$$

The final fraction of defaulted institutions under fire sales is determined by

$$f_D(z^\star(\bar{p});\bar{p}) = 1 - \sum_{\theta=1}^{d} q(\theta;\bar{p})\beta(d,z,d-\theta+1) - q(\infty;\bar{p}).$$

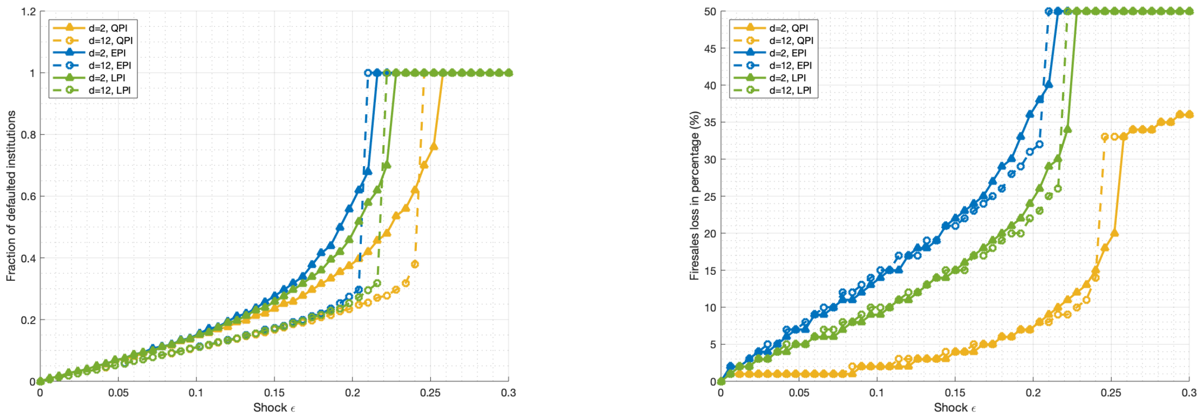
Furthermore, the final fraction of defaults without fire sales (with the initial price $p_0 = 2$) is

$$f_D(z^\star(2); 2) = 1 - \sum_{\theta=1}^{d} q(\theta; 2)\beta(d, z, d - \theta + 1) - q(\infty; 2).$$

In the financial network with low connectivity, we set the degree $d_L = 2$ and use the capital vector $\mathbf{h} = [50 \quad 100 \quad 250 \quad 300]$. For the network with high connectivity, we set the degree $d_H = 12$ and, for comparison purposes, use the same capital vector as the low connectivity network. To ensure the same total expected interbank liabilities, we introduce a dependency between the interbank liabilities and the degree. Specifically, the expectation of liabilities is assumed to be proportional to $1/d$. For the low connectivity network, we set $x_m = 160$ and $\alpha = 2$. Correspondingly, for the high connectivity network, we set $x_m = 26.7$ and $\alpha = 2$.

Figure 3.1a dispays the final fraction of defaulted institutions for the two regular financial networks with low and high connectivity, considering three price impact functions: linear (LPI) g^L, fully quadratic (QPI) g^Q_∞, and exponential (EPI) g^E_1. As expected, we observe that the EPI function leads to the largest fraction of defaults among the three price impact functions for both low and high connectivity networks. On the other hand, the QPI function results in the smallest default cascade size. This behavior arises because, for the same amount of sold shares, the EPI function always yields the lowest price, while the QPI function produces the highest price. Furthermore, we note the presence of a critical shock value (dependent on the connectivity and price impact function) where all institutions default. Interestingly, in the low connectivity network, the default cascade size increases smoothly as the shock magnitude increases. In contrast, the high connectivity network exhibits a sharper phase transition at the critical point. Additionally, when the shock is smaller than the critical value, the fraction of defaults increases slowly and remains lower than that in the low connectivity network. However, once the shock surpasses the critical value, the fraction of defaults jumps to a higher level than in the low connectivity network. This phenomenon aligns with existing literature on homogeneous financial networks, such as [1], which suggests that high connectivity networks are more resilient to small shocks but become more susceptible to large shocks due to their greater interconnections.

Figure 3.1b showcases the fire sales loss for the two regular networks with low and high connectivity, considering the three price impact functions. Since the fraction of defaults corresponds to the amount of liquidations, the curves in Figure 3.1b exhibit similar trends to those in Figure 3.1a. We observe that the EPI function consistently results in the largest fire sales loss, while the QPI function leads to the smallest fire sales loss. Interestingly, we also notice that the fire sales losses in the two networks with different connectivity levels are very close to each other. In fact, for small shocks (less than 0.15), the high connectivity network may even generate higher fire sales losses compared to the low connectivity network, which contrasts with the observations in Figure 3.1a. This occurs because in a higher connectivity network, institutions with high thresholds can still remain solvent while liquidating a significant portion (around 80-90%) of their total illiquid assets. On the other hand, in a lower connectivity network, the amount of liquidations among solvent institutions is much lower. As the shock increases but remains below the critical value, the fire sales loss in the low connectivity network may surpass that in the high connectivity network, regardless of the price impact function employed.



(a) Fraction of defaulted institutions (b) Fire sales loss

Figure 3.1: Final fraction of defaulted institutions and fire sales loss for two regular financial networks with $d_L = 2$ and $d_H = 12$, under three different price impact functions LPI, QPI and EPI.

3.4.3 Core-Periphery financial networks

Financial networks often exhibit significant asymmetries, such as the presence of a core-periphery structure. This structural characteristic has a notable impact on the size of default cascades. In such networks, large core institutions may be more resilient to small shocks compared to peripheral institutions. However, when core institutions experience a large shock, their default can trigger a substantial increase in the size of the default cascade. In our analysis, we do not impose a specific inter-structure for the core and peripheral banks but assume a random uniform connection between them. We consider two distinct classes of institutions, denoted as $\mathcal{X} = \{C, P\}$, representing the core institutions and peripheral institutions, respectively.

In our analysis, we assume a fraction of core and peripheral institutions, with $\mu_C = 0.3$ representing the core institutions and $\mu_P = 0.7$ representing the peripheral institutions. For the core institutions, we set the degree $d_C = 12$ with illiquid asset holdings $\gamma_C = 160$. For the peripheral type institutions, we set the degree $d_P = 2$ with $\gamma_P = 60$. Correspondingly, the capital structure vector for core institutions is set to $\mathbf{h}_C = [160 \quad 320 \quad 800 \quad 960]$, and for peripheral institutions, it is set to $\mathbf{h}_P = [60 \quad 120 \quad 300 \quad 360]$. As a result, the average degree of the financial network is given by $\lambda = 0.3 \times 12 + 0.7 \times 2 = 5$.

In our numerical experiments, we compare the core-periphery network described above to a 5-regular financial network with the same average degree. For the 5-regular network, we set the capital structure of all institutions to be the same as the average capital structure of the core-periphery network, denoted as $\bar{\mathbf{h}}$. Thus, the capital structure for each institution in the 5-regular network is set to $\bar{\mathbf{h}} = [90 \quad 180 \quad 450 \quad 540]$. To model the interbank liabilities in both networks, we assume that they are i.i.d. with a Pareto distribution, as in Example 3.4, with the scale and the shape parameters $x_m = 65$ and $\alpha = 2$, respectively.

Let q_C and q_P denote the probability threshold distributions for the core type and periphery type institutions, respectively. In this case, the limiting function f_W simplifies to

$$\begin{aligned} f_W(z;p) =& 5z - 3.6\big(\sum_{\theta=1}^{12} q_C(\theta;p)\beta(12,z,12-\theta+1) + q_C(\infty;p)\big) \\ & - 1.4\big(\sum_{\theta=1}^{2} q_P(\theta;p)\beta(2,z,2-\theta+1) + q_P(\infty;p)\big), \end{aligned}$$

and the limiting function for the total liquidations simplifies to

$$\begin{aligned} f_\Gamma(z;p) =& 48q_C(0;p) + \sum_{\theta=1}^{12} \frac{48}{p\theta} q_C(\theta;p)\big(\theta - \sum_{\ell=12-\theta+1}^{12} \beta(12,z,\ell)\big) + \frac{24}{p}(1-z)q_C(\infty;p) \\ & + 42q_P(0;p) + \sum_{\theta=1}^{2} \frac{42}{p\theta} q_P(\theta;p)\big(\theta - \sum_{\ell=2-\theta+1}^{2} \beta(2,z,\ell)\big) + \frac{21}{p}(1-z)q_C(\infty;p). \end{aligned}$$

Let $\bar{p}_{cp} = \bar{p}_{cp}(\epsilon)$ given by (3.13) be the limit for the price of illiquid asset in equilibrium after shock ϵ, as stated in Theorem 3.11. Then the limiting fire sales loss can be written as

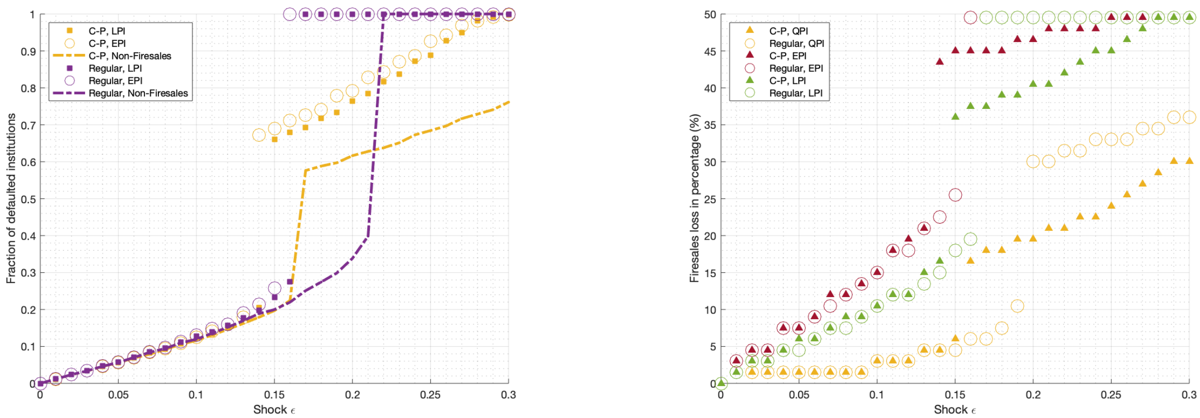
$$\mathrm{FSL}(\epsilon) = \frac{p_0 - \bar{p}_{cp}(\epsilon)}{p_0},$$

and the final fraction of defaulted institutions under fire sales is given by

$$\begin{aligned} f_D(z^\star(\bar{p}_{cp});\bar{p}_{cp}) =& 1 - 0.3\big(\sum_{\theta=1}^{12} q_C(\theta;\bar{p}_{cp})\beta(12,z,12-\theta+1) + q_C(\infty;\bar{p}_{cp})\big) \\ & - 0.7\big(\sum_{\theta=1}^{2} q_P(\theta;\bar{p}_{cp})\beta(2,z,2-\theta+1) + q_P(\infty;\bar{p}_{cp})\big). \end{aligned}$$

In Figure 3.2a, we plot the final fraction of defaulted institutions for the core-periphery network and compare it with the average regular network, in the cases without fire sales and with fire sales, considering the linear (LPI) and exponential (EPI) price impact functions. We observe that the fire sales make both networks more vulnerable. Without fire sales, the core-periphery network has a critical shock value around 0.16 (beyond which all institutions default), while for the regular network, the critical shock value is around 0.21. With fire sales, both financial networks have a smaller critical value for the shock. As expected, the EPI function gives a smaller critical value compared to the LPI function for both networks. Additionally, we note that the fire sales reduce the gap between the two critical shock values (for the core-periphery and regular networks). Without fire sales, the gap is about 0.05, but with fire sales (under both LPI and EPI functions), the gap is reduced to around 0.01. This can be interpreted by the fact that under fire sales, institutions have smaller thresholds θ to default, since $q_x(\theta;p)$ (stochastically) decreases with price p.

In Figure 3.2b, we plot the fire sales loss for the core-periphery network and compare it with the average regular network under different price impact functions (LPI, QPI, and EPI). We observe that



(a) Fraction of defaulted institutions (b) Fire sales loss

Figure 3.2: Final fraction of defaulted institutions and fire sales loss for core-periphery (C-P) and (average) regular financial networks, under three different price impact functions LPI, QPI and EPI.

under each price impact function, the regular and core-periphery networks perform very similarly when the shock is small (less than 0.15). However, when the shock is larger, the core-periphery network has less fire sales loss than the regular network. This occurs because a large portion of periphery institutions in the core-periphery network liquidate less than the average level. Additionally, the regular network has a larger critical shock value (beyond which all institutions default) compared to the core-periphery network. The smaller critical value for the core-periphery network is influenced by the core institutions, as their high degree makes them more likely to trigger a larger default cascade.

3.4.4 Scale-Free financial networks

Many empirically observed interbank networks have much more heterogeneity than the core-periphery financial network studied in the previous section. In order to study the effect of heterogeneity in network structure on the final size of default cascade and fire sales loss, we compare the following networks: Regular network (without heterogeneity), Erdös-Rényi random network (with low heterogeneity where the majority of institutions have a degree close to the average degree), and the Scale-free network (with high heterogeneity). To facilitate a meaningful comparison, we ensure that these networks have the same average degree λ.

For the Erdös-Rényi network, denoted by $\mathrm{ER}(n; p_n)$, each pair of nodes (a potential directed link) is independently connected with a fixed probability p_n such that $np_n \to \lambda$ as $n \to \infty$. In this network, the degree distribution converges to a Poisson distribution with parameter λ. That is, if we denote the in-degree or out-degree of a randomly chosen institution by D, then the probability mass function of D is given by

$$\mathbb{P}(D = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

On the other hand, for the scale-free network, the degree distribution follows a power law distribution, given by

$$\mathbb{P}(D = k) \sim ck^{-\eta},$$

where $c > 0$ is a normalizing constant and $\eta > 1$ is a control parameter.

We set the parameters $\lambda = 5$ and $\eta = 1.2$. To reduce the complexity of the simulation, we assume that the degrees are upper-bounded by $d_{\max} = 23$. These parameter choices result in both the scale-free and Erdös-Rényi networks having an average degree very close to 5. We will compare these networks to a regular network with a degree of 5.

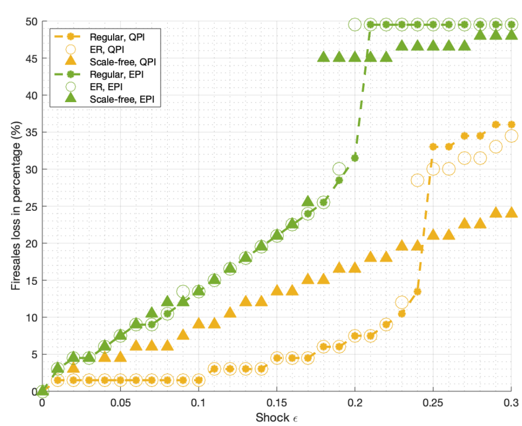
We also introduce heterogeneity in the interbank liabilities for all these networks. Specifically, we consider i.i.d. Pareto-distributed liabilities, as in Example 3.4, with scale and shape parameters $x_m = 55$ and $\alpha = 2$. Furthermore, we allow institutions with different degrees to have different capital structures, where the capital is proportional to the degree of each institution. For institutions with a degree of 1, we set the capital vector $\mathbf{h}_1$ as

$$\mathbf{h}_1 = [50 \quad 100 \quad 250 \quad 300].$$

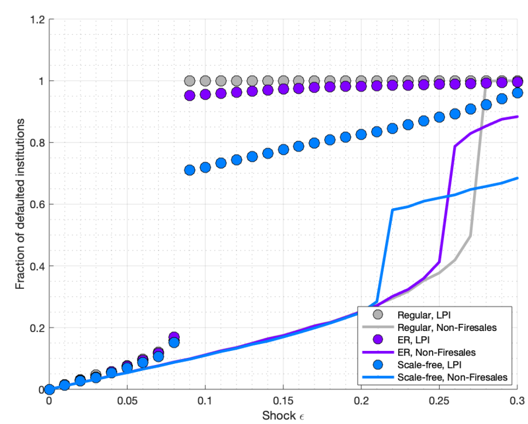
For degrees $d = 2, \ldots, 23$, we set the capital vector $\mathbf{h}_d$ as

$$\mathbf{h}_d = [10d + 40 \quad 20d + 80 \quad 50d + 200 \quad 60d + 240].$$

In the case of the regular network, where all institutions have a degree of 5, the capital structure for each institution is given by $\mathbf{h}_5 = [90 \quad 180 \quad 450 \quad 540]$, as in the previous section.



(a) Fire sales loss under the QPI and EPI functions



(b) Final fraction of defaults under the LPI and without fire sales

Figure 3.3: Fire sales loss and final faction of defaults for regular, Erdös-Rényi (ER) and scale-free networks.

In Figure 3.3a, we compare the fire sales loss for the regular, Erdös-Rényi (ER) , and scale-free networks under the quadratic price impact (QPI) and exponential price impact (EPI) functions. We observe that for the EPI function, when the shock is small (less than 0.17, which is the critical shock value for the scale-free network), the heterogeneity does not have a significant influence on the fire sales loss. However, for the QPI function, when the shock is small, the fire sales loss in the scale-free network is larger than the fire sales loss in the other two networks. For a small shock (less than 0.1), the fire sales loss is only about 0.2% for both the ER and regular networks. This difference occurs because choosing an ϵ-fraction of initially defaulted institutions at random among all institutions can lead to a small fraction of initial defaults for high-degree institutions, which in turn can result in a considerable fraction of defaults among low-degree institutions, leading to more fire sales loss. This effect is particularly significant under the slow-dropping price impact function.

Moreover, as we can observe in Figure 3.3a, a network with higher heterogeneity has a smaller critical value for the shock (beyond which a large fraction of institutions default). When the shock is larger than the critical value for the regular network (around 0.2 under EPI and 0.24 under QPI), the regular network has the largest fire sales loss, while the scale-free network has the smallest loss. This is reasonable because in the scale-free network, there is a larger proportion of institutions with low degrees (such as 1 and 2), which have a higher chance of surviving for a large value of shock. This makes the scale-free network more resilient to a large shock compared to the other two networks.

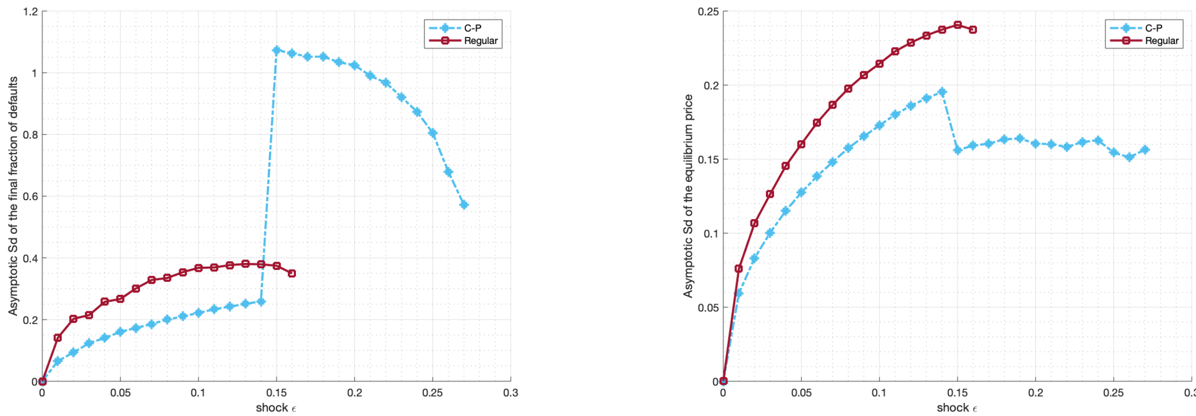
Figure 3.3b displays the final fraction of defaulted institutions for the regular, Erdös-Rényi (ER), and scale-free networks, for the case without fire sales and with the linear price impact (LPI) fire sales. We can observe similar results as in Figure 3.2a. When the shock is small, the fire sales do not have a significant impact on the resilience of the networks. However, the fire sales significantly reduce the critical values for shocks in all three networks. The critical values become closer to each other, with the regular, ER, and scale-free networks having critical values around 0.11 under the linear price impact function. Among the three networks, the scale-free network has the smallest critical value for the shock, followed by the ER network, while the regular network has the largest critical value. Moreover, the resistance to a large shock increases with heterogeneity, especially under the fire sales impact. The scale-free network has the smallest fraction of defaults for a shock larger than 0.1. Therefore, as observed from Figure 3.2a and Figure 3.3b, networks with higher heterogeneity tend to have smaller critical values for shocks, beyond which a large fraction of institutions default, both with and without fire sales. On the other hand, for small shocks, the most heterogeneous network could be the least resilient.

3.4.5 Asymptotic normality and confidence intervals

To demonstrate the practical relevance of central limit theorems, we analyze the asymptotic variance for the final fraction of defaults and fire sales loss. Additionally, we utilize central limit theorems to derive confidence intervals for financial networks of finite size.

It is worth noting that the scale-free network discussed in Section 3.4.4, which exhibits infinite variance in the degree distribution, does not satisfy Assumption 3.5b. As a result, we focus our analysis on a regular network and a core-periphery network, both having an average degree of 5 as shown in Figure 3.2. All parameters remain the same as in Section 3.4.3, and we specifically consider the linear price impact function. For each institution $x \in \mathcal{X}$ and threshold $\theta = 1, \ldots, d_x^+$, we set the variance of each liquidation $\varsigma_{x,\theta}^2(p)$ equal to the mean of each liquidation, that is, $\varsigma_{x,\theta}^2(p) = \bar{\ell}_{x,\theta}(p)$ for all $x \in \mathcal{X}$,

$\theta \in \{1, \ldots, d_x\} \cup \{\infty\}$, and all $p \in [1, 2]$.



(a) Fraction of defaulted institutions

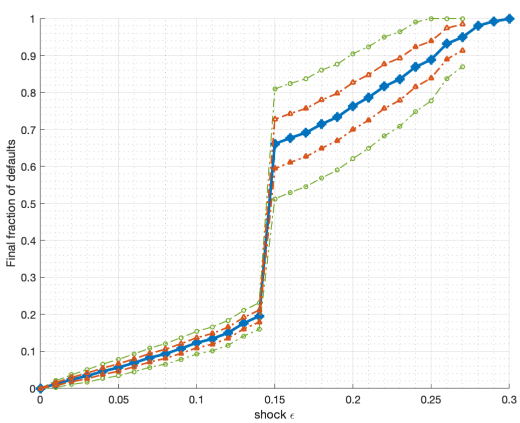
(b) Fire sales loss

Figure 3.4: Asymptotic standard deviation for the final fraction of defaulted institutions and fire sales loss in regular and core-periphery financial networks, under linear price impact function (LPI).

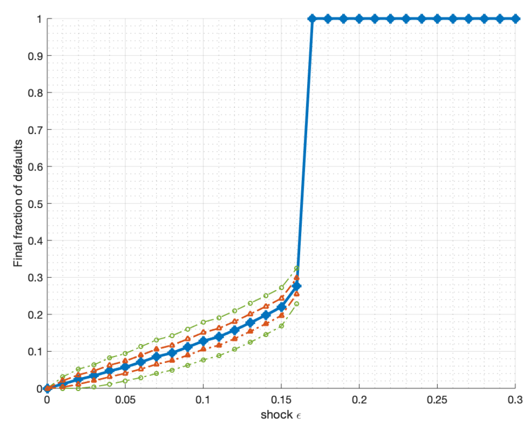
In Figure 3.4a, we present the asymptotic standard deviation (Sd) of the final fraction of defaulted institutions under different shocks. The blue line represents the standard deviation for the core-periphery network, while the red line corresponds to the standard deviation for the regular network. Since our central limit theorems hold when $z^\star \neq 0$, we terminate the standard deviation plots at the point where $z^\star(\epsilon) = 0$, which corresponds to the critical value of the shock where almost all institutions default ($\epsilon = 0.16$ for the regular network and $\epsilon = 0.27$ for the core-periphery network). We observe that for small shocks, the standard deviation increases with the shock, indicating that a larger number of defaults leads to higher uncertainty in the final fraction of defaulted institutions. Furthermore, for small shocks, the standard deviation of the fraction of defaults is larger for the regular network compared to the core-periphery network. This can be attributed to the fact that for small shocks, most of the defaults occur among periphery institutions with a degree of 2, resulting in less variability compared to the regular network. Surprisingly, we observe a significant jump in the standard deviation at around $\epsilon \approx 0.14$, followed by a decrease with further increases in the shock. This critical value aligns with the point of discontinuity in the fixed point solution, where the fraction of defaults jumps to a higher level (but still smaller than 1) for the core-periphery network. In contrast, for the regular network, the fraction of defaults jumps to 1. This observation is consistent with the findings in Figure 3.2a, which demonstrate the sharp phase transition in the core-periphery network compared to the smoother increase in the regular network.

Figure 3.4a displays the asymptotic standard deviation (Sd) of final fraction of defaulted institutions under different shocks. The blue line is the standard deviation for the core-periphery network and the red line is the standard deviation for the regular network. Since our central limit theorems hold when $z^\star \neq 0$, we stop the standard deviation plots at the point when $z^\star(\epsilon) = 0$, which corresponds to the critical value for shock such that almost all institutions default ($\epsilon = 0.16$ and $\epsilon = 0.27$ for the regular and core-periphery networks, respectively). We observe that when the shock is small, the standard deviation for both networks are increasing with shock and more defaults cause larger standard deviation. Moreover, for small shocks, the standard deviation of fraction of defaults for regular network is larger than that of core-periphery network. This can be justified by the fact that, for small shocks, most defaults are periphery institutions, with small degree 2, which make less uncertainty compare to the regular network. Surprisingly, the standard deviation exhibits a significant jump at around $\epsilon \approx 0.14$ and then starts to decrease with increasing shock. As Figure 3.2a shows, this critical value corresponds to the point where a discontinuity occurs at the fixed point solution, and the fraction of defaults jumps to a higher level; in the core-periphery network, the fraction of defaults does not reach 1 at this critical point, unlike the regular network where the fraction of defaults jumps to 1.

In Figure 3.4b, we plot the asymptotic standard deviation of fire sales loss under different shocks. As expected, the shapes of the curves exhibit similarities to those for the standard deviation of the fraction of defaults. However, the standard deviation of fire sales loss for the core-periphery financial network displays a downward discontinuity jump. This can be explained by considering the standard deviation of the total amount of liquidations, which is influenced by both the standard deviation and the mean number of liquidations. It is important to note that the fraction of defaults provides only partial information on the number of liquidations, as some institutions may remain solvent despite having already liquidated a portion of their holdings. After the discontinuity jump in the fraction of defaults, the standard deviation of the number of liquidations decreases and may exhibit a downward discontinuity jump as well, as most institutions have already liquidated all their holdings.



(a) Core-periphery network

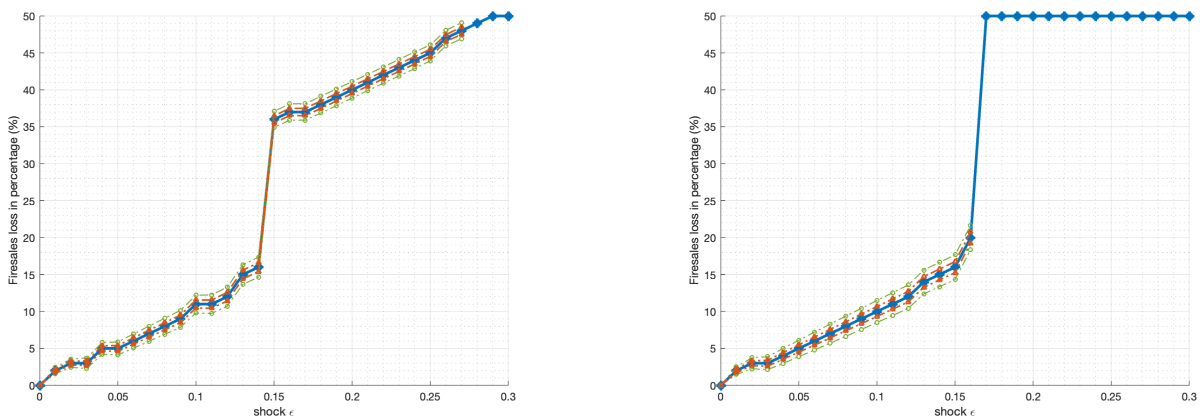


(b) Regular network

Figure 3.5: 95% confidence intervals for the final fraction of defaulted institutions in regular and core-periphery financial networks. The blue solid line is the limit; the green dash lines are the upper and lower bounds in the case of network size $n = 200$; the red dash lines are the bounds for network size $n = 1000$.

Our central limit theorems can be used to provide confidence intervals for the final fraction of defaults and fire sales loss in finite networks. In Figure 3.5, we plot the 95% confidence interval for the final fraction of defaults for both regular and core-periphery financial networks under different shocks. In both networks, when the shock is smaller than the critical value, the confidence intervals are quite small. Even for a small network size of $n = 200$, the maximum distance from the upper or lower bounds to the limit value (the blue solid line) is approximately 0.05. As the network size increases to $n = 1000$, the confidence intervals remain uniformly small for both regular and core-periphery financial networks.

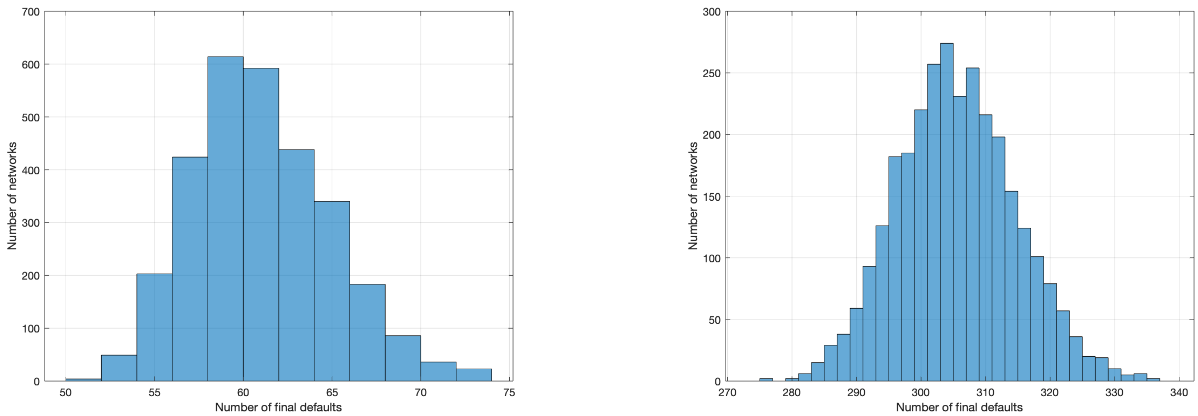
Figure 3.6 displays the 95% confidence interval for the fire sales loss in both regular and core-periphery networks under different shocks. In the core-periphery network, the confidence intervals are concentrated closely around the limits. As shown in Figure 3.6a and Figure 3.6b, even for a small network size of $n = 200$, the upper and lower bounds are very close to the limits across the entire shock range. When $n = 1000$, the upper and lower bounds almost coincide with the limits for both regular and core-periphery networks.



(a) Core-periphery network (b) Regular network

Figure 3.6: 95% confidence intervals for the fire sales loss in regular and core-periphery financial networks. The blue solid line is the limit, the green dash lines are the bounds for network size $n = 200$, and the red dash lines are the bounds for network size $n = 1000$.

To study the convergence of our central limit theorems numerically, we consider networks with finite size n and simulate the final fraction of defaulted institutions using the Monte Carlo method. For simplicity, we fix the price at $p = 2$ and simulate only regular networks. In order to observe the convergence of the distribution of the final fraction of defaults to a Gaussian distribution as n becomes large, we run 3000 simulations of the default cascade in different regular networks with a degree of $d = 5$ chosen uniformly at random from all 5-regular networks. We count the number of institutions that default at the end of each simulation and produce histograms based on these counts. Figure 3.7 displays the histograms obtained for two different network sizes, $n = 500$ and $n = 2500$. As we can see, when $n = 2500$, the distribution already closely resembles a Gaussian distribution.



(a) Network size n=500 (b) Network size n=2500

Figure 3.7: Histograms for the final number of defaulted institutions in 5-regular financial networks with size $n = 500$ and $n = 2500$, using 3000 times Monte-Carlo simulations.

3.5 Proofs

This section includes the proofs of all theorems from the previous sections. The work is a generalization of Chapter 2 with significantly more involved computations. In Chapter 2, we investigate specific limit theorems connected to the final state of the pure default cascade death process from Section 3.2.2, without taking into account the price p of the illiquid asset. Here, we use these results and adapt them to fit our framework that includes price p, without repeating the proofs. Importantly, for each fixed $p \in [p_{\min}, p_0]$, these results still apply. The challenge arises as the liquidations are type and threshold-dependent i.i.d. random variables. This makes the total liquidation turn into a compound (random) sum of random variables. These random sums are governed by counting stochastic processes that exhibit Gaussian fluctuations in the limit. Under these conditions, proving the central limit theorems for the total liquidation of the network becomes more complicated. After obtaining the limit results for the liquidation, the price given by the inverse demand function can be found using the Delta method. Lastly, for the equilibrium price - as a solution of an equation - the limit theorems are set under stronger assumptions, demanding a more complex technical proof. Compared to Chapter 2, we face more variation and handle more difficult convergence problems. We start the proofs by providing some auxiliary lemmas used in the proof of the central limit theorems.

3.5.1 Auxiliary lemmas

Under certain regularity conditions, we first provide a central limit theorem for functions that can be expressed as $Y_n(t) := \sum_{i=1}^{\lfloor X_n(t) \rfloor} G_i$, where $X_n(t)$ is a non-decreasing stochastic process satisfying $X_n(t) = O(n)$ for all $t > 0$ and $\{G_i\}_{i \geqslant 1}$ are i.i.d. positive bounded random variables with mean g and variance σ^2.

Lemma 3.20. *Using the notation above and for fixed $t > 0$, if $X_n(t) := f_n(t)n + \mathcal{V}n^{1/2}$ with $(f_n(t))_{n=1}^{\infty}$ a positive sequence converging to $f(t)$, and $\mathcal{V}$ a bounded real-valued random variable, then as $n \to \infty$, conditioned on $\{\mathcal{V} = x\}$ for some x on supp($\mathcal{V}$), we have*

$$\Big(\frac{Y_n(t) - gX_n(t)}{\sqrt{nf(t)}\sigma} \mid \mathcal{V} = x\Big) \xrightarrow{d} \mathcal{N}(0,1).$$

Proof. Conditioned on the event $\{\mathcal{V} = x\}$, $X_n(t) = f_n(t)n + xn^{1/2}$ which is non-random. Hence, by standard central limit theorem (CLT), we have

$$\Big(\frac{Y_n(t) - g\lfloor X_n(t)\rfloor}{\sqrt{\lfloor nf_n(t) + xn^{1/2}\rfloor}\sigma}|\mathcal{V} = x\Big) \xrightarrow{d} \mathcal{N}(0,1).$$

Further, we have the decomposition

$$\begin{aligned}\frac{Y_n(t) - gX_n(t)}{\sqrt{nf(t)}\sigma} &= \frac{\sqrt{\lfloor nf_n(t) + xn^{1/2}\rfloor}}{\sqrt{nf(t)}} \cdot \frac{Y_n(t) - g\lfloor X_n(t)\rfloor}{\sqrt{\lfloor nf_n(t) + xn^{1/2}\rfloor}\sigma} + \frac{g\lfloor X_n(t)\rfloor - gX_n(t)}{\sqrt{nf(t)}\sigma} \\ &= \sqrt{1 + O(n^{-1/2})}\frac{Y_n(t) - g\lfloor X_n(t)\rfloor}{\sqrt{\lfloor nf(t) + xn^{1/2}\rfloor}\sigma} + O(n^{-1/2}).\end{aligned}$$

It follows thus by Slutsky's theorem that as $n \to \infty$,

$$\Big(\frac{Y_n(t) - gX_n(t)}{\sqrt{nf(t)}\sigma}|\mathcal{V} = x\Big) \xrightarrow{d} \mathcal{N}(0,1).$$

□

Using the above lemma, we prove the following proposition.

Proposition 3.21. *For fixed $t > 0$, let $X_n(t) := f_n(t)n + \mathcal{V}_n n^{1/2}$ with $\{f_n(t)\}_{n=1}^{\infty}$ a positive sequence converging to $f(t)$ and $\mathcal{V}_n$ a sequence of random variables which converges to a Gaussian random variable $\mathcal{V} \sim \mathcal{N}(0, v^2)$ in distribution. Then we have, as $n \to \infty$,*

$$\frac{Y_n(t) - ngf_n(t)}{\sqrt{n(f(t)\sigma^2 + v^2g^2)}} \xrightarrow{d} \mathcal{N}(0,1).$$

Proof. Set

$$A(z;p) := \int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{f(t)}\sigma r} \exp\left\{-\frac{u^2}{2v^2} - \frac{(z-gu)^2}{2f(t)\sigma^2}\right\} du. \tag{3.18}$$

Let $a := v^2g^2 + f(t)\sigma^2$. Then by a change of variable $y = \frac{\sqrt{a}}{r\sigma\sqrt{f(t)}}u - \frac{rgz}{\sigma\sqrt{af(t)}}$, we obtain

$$\begin{aligned}A(z;p) &= \frac{1}{2\pi\sqrt{f(t)}\sigma r}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2f(t)\sigma^2v^2}(f(t)\sigma^2u^2 + z^2v^2 - 2gzuv^2 + v^2g^2u^2)\right\} du \\ &= \frac{1}{2\pi\sqrt{f(t)}\sigma r}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2f(t)\sigma^2v^2}((\sqrt{a}u - \frac{v^2gz}{\sqrt{a}})^2 + \frac{f(t)\sigma^2v^2z^2}{a})\right\} du \\ &= \frac{1}{2\pi\sqrt{f(t)}\sigma r}e^{-\frac{z^2}{2a}}\int_{-\infty}^{\infty} \frac{r\sigma\sqrt{f(t)}}{\sqrt{a}}e^{-\frac{y^2}{2}}dy = \frac{1}{\sqrt{2\pi a}}e^{-\frac{z^2}{2a}}.\end{aligned}$$

We consider the function

$$h_z(x) := \frac{1}{\sqrt{2\pi f(t)\sigma}} \exp\left\{-\frac{(z-gx)^2}{2f(t)\sigma^2}\right\},$$

which is continuous and bounded. Since $\mathcal{V}_n \xrightarrow{d} \mathcal{V}$, we have (as $n \to \infty$)

$$A_n(z;p) := \mathbb{E}[h_z(\mathcal{V}_n)] \longrightarrow \mathbb{E}[h_z(\mathcal{V})] = A(z;p).$$

We denote

$$Z_n(t) := \frac{Y_n(t) - ngf_n(t)}{\sqrt{n}},$$

and let $Z(t)$ be a random variable with distribution $Z(t) \sim \mathcal{N}(0, \sigma^2 f(t))$. Let μ_n be the probability measure of $\mathcal{V}_n$ and μ be that of $\mathcal{V}$. Set

$$\Phi_x(B) := \mathbb{P}(Z(t) - gx \in B),$$

and

$$G_{\mathcal{V}_n}(B|x) := \mathbb{P}(Z_n(t) \in B | \mathcal{V}_n = x).$$

Then for any Borel set $B \subset \mathbb{R}$, we have

$$\begin{aligned}
\big|\mathbb{P}(Z(t) \in B) - \mathbb{P}(Z_n(t) \in B)\big| =& \Big|\int_{\mathbb{R}} \Phi_x(B) d\mu(x) - \int_{\mathbb{R}} G_{\mathcal{V}_n}(B|x) d\mu_n(x)\Big| \\
\leqslant& 2\epsilon + \Big|\int_{[-K,K]} G_{\mathcal{V}_n}(B|x) d\mu_n(x) - \int_{[-K,K]} \Phi_x(B) d\mu_n(x)\Big| \\
&+ \Big|\int_{\mathbb{R}} \Phi_x(B) d\mu_n(x) - \int_{\mathbb{R}} \Phi_x(B) d\mu(x)\Big|,
\end{aligned}$$

where we take K large enough such that $\int_{\mathbb{R}\setminus[-K,K]} 1 d\mu_n(x) \leqslant \epsilon$, uniformly on n.

We have

$$\Big|\int_{[-K,K]} G_{\mathcal{V}_n}(B|x) d\mu_n(x) - \int_{[-K,K]} \Phi_x(B) d\mu_n(x)\Big| \to 0.$$

Indeed, since any Borel set is a continuity set of Gaussian distribution, for every $x \in \operatorname{supp}(\mathcal{V}_n) \cap [-K, K]$, $G_{\mathcal{V}_n}(B|x) \to \Phi_x(B)$ by Lemma 3.20. The result follows by the dominant convergence theorem. Moreover, we have

$$\begin{aligned}
&\Big|\int_{\mathbb{R}} \Phi_x(B) d\mu_n(x) - \int_{\mathbb{R}} \Phi_x(B) d\mu(x)\Big| \leqslant \int_B |\mathbb{E}[h_z(\mathcal{V}_n)] - \mathbb{E}[h_z(\mathcal{V})]| dz \\
\leqslant& \int_{\mathbb{R}} |A_n(z;p) - A(z;p)| dz \to 0,
\end{aligned}$$

where the first inequality follows by Fubini's theorem and the second by Scheffé's lemma since $\int_{\mathbb{R}} A_n(z;p)dz = \int_{\mathbb{R}} A(z;p)dz = 1$ and $A_n(z;p) \to A(z;p)$ for every $z \in \mathbb{R}$.

Since we can choose ϵ arbitrarily, we finally get for any borel set $B \in \mathbb{R}$,

$$\mathbb{P}(Z_n(t) \in B) \to \int_B A(z;p) dz.$$

Since $A(z;p)$ is the density of $\mathcal{N}(0,a)$ and all Borel sets are continuity set of $\mathcal{N}(0,a)$, it follows that $Z_n(t) \xrightarrow{d} \mathcal{N}(0,a)$, which is equivalent to the statement of proposition. □

3.5.2 Proof of Theorem 3.7

By Lemma 3.6 for all $x \in \mathcal{X}, \theta = 1, \dots, d_x^+, \ell = 0, \dots, \theta - 1$ and $p \in [p_0, p_{\min}]$, as $n \to \infty$,

$$\sup_{t \leqslant \tau_n} \Big| \frac{S_{x,\theta,\ell}^{(n)}(t;p)}{n} - \mu_x q_x(\theta;p) b\left(d_x^+, 1 - e^{-t}, \ell\right) \Big| \xrightarrow{p} 0,$$

and,

$$\sup_{t \leqslant \tau_n} \Big| \frac{S_n(t;p)}{n} - f_S(e^{-t};p) \Big| \xrightarrow{p} 0, \quad \sup_{t \leqslant \tau_n} \Big| \frac{D_n(t;p)}{n} - f_D(e^{-t};p) \Big| \xrightarrow{p} 0.$$

Consider the death process as described in Chapter 2. We denote by $U_{x,\theta,s}^{(n)}(t;p)$ the number of institutions with type $x \in \mathcal{X}$, threshold θ and s alive incoming half-edges at time t, and by $N_{x,\theta}^{(n)}(p)$ the number of institutions with type x and threshold θ, under price p. Note that the number of institutions with type x is (not random) $n\mu_x^{(n)}$. By construction of the death process model, each incoming hal-edge has an exponentially distributed with parameter one, i.e., exp(1), lifetime independently from others. Using the Glivenko-Cantelli theorem, in Lemma 2.10 in Chapter 2 we show the following convergence results of $U_{x,\theta,\ell}^{(n)}(t;p)$, for all possible triple (x, θ, ℓ) and the summation of them. To make the proof clear, we state it again and adapt to the price dependent case.

Lemma 3.22. *Let $\tau_n \leqslant \tau_n^\star(p)$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Under Assumption 3.5a, for all $x \in \mathcal{X}, \theta = 1, \dots, d_x^+$ and $\ell = 0, \dots, \theta - 1$, we have (as $n \to \infty$)*

$$\sup_{t \leqslant \tau_n} \Big| \frac{U_{x,\theta,\ell}^{(n)}(t;p)}{n} - \mu_x q_x(\theta;p) b\left(d_x^+, e^{-t}, \ell\right) \Big| \xrightarrow{p} 0.$$

Further,

$$\sup_{t \leqslant \tau_n} \sum_{x \in \mathcal{X}} (d_x^+ + d_x^-) \sum_{\theta=1}^{d_x^+} \sum_{s=d_x^+-\theta+1}^{d_x^+} \left| U_{x,\theta,s}^{(n)}(t;p)/n - \mu_x q_x(\theta;p) b(d_x^+, e^{-t}, s) \right| \xrightarrow{p} 0.$$

Consider now $\{L_{x,\theta}^{(i)}(p)\}_{i=1}^\infty$ which are i.i.d. positive bounded random variables with expectation $\bar{\ell}_{x,\theta}(p)$ and variance $\varsigma_{x,\theta}^2(p)$ under price $p \in [p_{\min}, p_0]$ for the illiquid asset, for all $x \in \mathcal{X}$ and $\theta \in \{1, \dots, d_x^+\} \cup \{\infty\}$. Since all the random losses are assumed to be bounded, we denote by C the common upper bound. From Section 3.2.4, the total shares of illiquid asset sold by time t can be written as

$$\Gamma_n(t;p) := \sum_{x \in \mathcal{X}} \Big(\bar{\gamma}_x D_{x,0}^{(n)}(p) + \sum_{\theta=1}^{d_x^+} Y_{x,\theta}^{(n)}(t;p), + Y_{x,\infty}^{(n)}(t;p) \Big),$$

where

$$Y_{x,\theta}^{(n)}(t;p) := \sum_{i=1}^{I_{x,\theta}^{(n)}(t;p)} L_{x,\theta}^{(i)}(p), \quad Y_{x,\infty}^{(n)}(t;p) := \sum_{i=1}^{I_{x,\infty}^{(n)}(t;p)} L_{x,\infty}^{(i)}(p),$$

and,

$$I_{x,\theta}^{(n)}(t;p) := \theta D_{x,\theta}^{(n)}(t;p) + \sum_{\ell=1}^{\theta-1} \ell S_{x,\theta,\ell}^{(n)}(t;p), \quad I_{x,\infty}^{(n)}(t;p) := \sum_{\ell=1}^{d_x^+} \ell S_{x,\infty,\ell}^{(n)}(t;p).$$

By Assumption 3.3 and the dominated convergence theorem, the first term in $\Gamma_n(t;p)$ converges to

$$\sum_{x\in\mathcal{X}} \bar{\gamma}_x D_{x,0}^{(n)}(p) \xrightarrow{p} \sum_{x\in\mathcal{X}} \bar{\gamma}_x \mu_x q_x(0;p).$$

Note that by definition $S_{x,\theta,\ell}^{(n)}(t;p) = U_{x,\theta,d_x^+-\ell}^{(n)}(t;p)$, which implies that

$$\sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta} \sum_{\ell=1}^{\theta-1} \ell S_{x,\theta,\ell}^{(n)}(t;p) = \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta} \sum_{s=d_x^+-\theta+1}^{d_x^+} (d_x^+ - s) U_{x,\theta,s}^{(n)}(t;p),$$

and,

$$\sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \theta \bar{\ell}_{x,\theta} D_{x,\theta}^{(n)}(t;p) = \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \theta \bar{\ell}_{x,\theta} \big(N_{x,\theta}^{(n)}(p) - \sum_{s=d_x^+-\theta+1}^{d_x^+} U_{x,\theta,s}^{(n)}(t;p) \big).$$

So for $\theta = 1, \ldots, d_x^+$, we have

$$\sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta} I_{x,\theta}^{(n)}(t;p) = \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \theta \bar{\ell}_{x,\theta} N_{x,\theta}^{(n)}(p) - \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta} \sum_{s=d_x^+-\theta+1}^{d_x^+} (s - d_x^+ + \theta) U_{x,\theta,s}^{(n)}(t;p).$$

Notice now that

$$\sum_{s=d_x^+-\theta+1}^{d_x^+} \beta(d_x^+, e^{-t}, s) = \sum_{s=d_x^+-\theta+1}^{d_x^+} (s - d_x^+ + \theta) b(d_x^+, e^{-t}, s).$$

From the definition $f_{x,\theta}(z;p) := \mu_x q_x(\theta;p)\big(\theta - \sum_{\ell=d_x^+-\theta+1}^{d_x^+} \beta(d_x^+, z, \ell)\big)$, it follows that

$$\begin{aligned}
&\Big|\frac{1}{n} \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} Y_{x,\theta}^{(n)}(t;p) - \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta} f_{x,\theta}(e^{-t};p)\Big| \\
&\quad \leqslant \Big| \sum_{x\in\mathcal{X}} d_x^+ \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta} \sum_{s=d_x^+-\theta+1}^{d_x^+} \big(U_{x,\theta,s}^{(n)}(t;p)/n - \mu_x q_x(\theta;p) b(d_x^+, e^{-t}, s) \big) \Big| \\
&\qquad + \frac{1}{n} \Big| \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} (Y_{x,\theta}^{(n)}(t;p) - \bar{\ell}_{x,\theta} I_{x,\theta}^{(n)}(t;p)) \Big| + \Big| \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \theta \bar{\ell}_{x,\theta} N_{x,\theta}^{(n)}(p)/n - \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \theta \bar{\ell}_{x,\theta} \mu_x q_x(\theta;p) \Big| \\
&\quad \leqslant C \sum_{x\in\mathcal{X}} (d_x^+ + d_x^-) \sum_{\theta=1}^{d_x^+} \sum_{s=d_x^+-\theta+1}^{d_x^+} \big| U_{x,\theta,s}^{(n)}(t;p)/n - \mu_x q_x(\theta;p) b(d_x^+, e^{-t}, s) \big| \\
&\qquad + \frac{1}{n} \Big| \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} (Y_{x,\theta}^{(n)}(t;p) - \bar{\ell}_{x,\theta} I_{x,\theta}^{(n)}(t;p)) \Big| + C \Big| \sum_{x\in\mathcal{X}} d_x^+ \sum_{\theta=1}^{d_x^+} \big(N_{x,\theta}^{(n)}(p)/n - \mu_x q_x(\theta;p) \big) \Big|.
\end{aligned}$$

The first term of the r.h.s. of the above inequality converges to 0, as $n \to \infty$, by Lemma 3.22. For the second term, note that for all $n, x \in \mathcal{X}$ and $t \leqslant \tau_n$, $\sum_{\theta=1}^{d_x^+}\big(Y_{x,\theta}^{(n)}(t;p) - \bar{\ell}_{x,\theta} I_{x,\theta}^{(n)}(t;p)\big)$ is a martingale. Combining this with the independency between any two different classes in $\mathcal{X}$, using Cauchy-Schwarz inequality and Doob's L^2-inequality, we have that for some constant C_0, as $n \to \infty$,

$$\mathbb{E}\big[\sup_{t\leqslant\tau_n}\frac{1}{n}\big|\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}(Y_{x,\theta}^{(n)}(t;p) - \bar{\ell}_{x,\theta} I_{x,\theta}^{(n)}(t;p))\big|\big]^2 \leqslant \frac{4C_0^2}{n^2}\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+} I_{x,\theta}^{(n)}(\tau_n) \leqslant \frac{4C_0^2}{n}\sum_{x\in\mathcal{X}}\mu_x^{(n)} d_x^+ \to 0,$$

where the second inequality above follows from $\sum_{\theta=1}^{d_x^+} I_{x,\theta}^{(n)}(t;p)/n \leqslant \mu_x^{(n)} d_x^+$ for all $t \leqslant \tau_n$. The final convergence holds by Assumption 3.5a. We next analyze the convergence result for the third term. First notice that, by the law of large numbers and Assumption 3.3, $N_{x,\theta}^{(n)}(p)/n \xrightarrow{p} \mu_x q_x(\theta;p)$. Let $\mathcal{X}_K^+$ be the set of all characteristic $x \in \mathcal{X}$ such that $d_x^+ \geqslant K$. Since by Assumption 3.5a, $\lambda \in (0,\infty)$, for arbitrary small $\varepsilon > 0$, there exists K_ε such that $\sum_{x\in\mathcal{X}_{K_\varepsilon}}\mu_x d_x^+ < \varepsilon$. Then by dominated convergence theorem, we obtain for n large enough,

$$\sum_{x\in\mathcal{X}_{K_\varepsilon}} d_x^+ \sum_{\theta=1}^{d_x^+} N_{x,\theta}^{(n)}(p)/n \xrightarrow{p} \sum_{x\in\mathcal{X}_{K_\varepsilon}} d_x^+ \sum_{\theta=1}^{d_x^+}\mu_x q_x(\theta) \leqslant \sum_{x\in\mathcal{X}_{K_\varepsilon}} d_x^+\mu_x < \varepsilon..$$

It follows that

$$C\big|\sum_{x\in\mathcal{X}} d_x^+ \sum_{\theta=1}^{d_x^+}\big(N_{x,\theta}^{(n)}(p)/n - \mu_x q_x(\theta;p)\big)\big| \leqslant C\sum_{x\in\mathcal{X}\setminus\mathcal{X}_{K_\varepsilon}} d_x^+\sum_{\theta=1}^{d_x^+}\big|N_{x,\theta}^{(n)}(p)/n - \mu_x q_x(\theta;p)\big| + C\varepsilon = o_p(1) + C\epsilon.$$

We conclude that

$$\sup_{t\leqslant\tau_n}\big|\frac{1}{n}\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+} Y_{x,\theta}^{(n)}(t;p) - \sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\bar{\ell}_{x,\theta} f_{x,\theta}(e^{-t};p)\big| \xrightarrow{p} 0.$$

It remains to prove the convergence for the third term (infinite sum) in $\Gamma_n(t;p)$. First notice that

$$I_{x,\infty}^{(n)}(t;p) = \sum_{s=0}^{d_x^+}(d_x^+ - s)U_{x,\infty,s}^{(n)}(t;p).$$

By using Lemma 3.22, for any type $x \in \mathcal{X}$, we have that

$$\sup_{t\leqslant\tau_n}\big|I_{x,\infty}^{(n)}(t;p)/n - \sum_{s=1}^{d_x^+}(d_x^+ - s)\mu_x q_x(\infty;p)b(d_x^+, e^{-t}, s)\big| \xrightarrow{p} 0.$$

Moreover,

$$\sum_{s=1}^{d_x^+}(d_x^+ - s)\mu_x q_x(\infty)b(d_x^+, e^{-t}, s) = \mu_x q_x(\infty)(d_x^+ - zd_x^+).$$

Then, by following a similar argument as above, one can show that

$$\sup_{t\leqslant\tau_n}\big|\frac{1}{n}\sum_{x\in\mathcal{X}} Y_{x,\infty}^{(n)}(t;p) - \sum_{x\in\mathcal{X}}\bar{\ell}_{x,\infty} f_{x,\infty}(e^{-t};p)\big| \xrightarrow{p} 0.$$

Putting all these convergence results together, we conclude that

$$\sup_{t \leqslant \tau_n} \left| \frac{\Gamma_n(t;p)}{n} - f_\Gamma(e^{-t};p) \right| \xrightarrow{p} 0,$$

as desired. □

3.5.3 Proof of Theorem 3.9

Fix $p \in [p_{\min}, p_0]$. The theorem follows from Theorem 3.7 and Lemma 3.8. Indeed, for $z^\star(p) = 0$, by Lemma 3.8, $\tau_n^\star(p) \xrightarrow{p} \infty$. Note that $z^\star(p) = 0$ indicates that almost all institutions default during the cascade. In this case, for all $x \in \mathcal{X}$, we have $q_x(\infty;p) = 0$. Otherwise $z^\star(p)$ can not be 0, since if $q_x(\infty;p) > 0$ for some $x \in \mathcal{X}$, then $f_W(0;p) < 0$. So $e^{-\tau_n^\star} \xrightarrow{p} 0$, and we have

$$f_\Gamma(0;p) = \sum_{x \in \mathcal{X}} \mu_x \Big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p) \theta q_x(\theta;p) \Big).$$

It follows by the continuity of f_Γ that

$$f_\Gamma(e^{-\tau_n^\star(p)};p) = \sum_{x \in \mathcal{X}} \mu_x \Big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p) \theta q_x(\theta;p) \Big) + o_p(1).$$

We therefore have by Theorem 3.7 that

$$\frac{\Gamma_n(\tau_n^\star(p))}{n} \xrightarrow{p} \sum_{x \in \mathcal{X}} \mu_x \Big(\bar{\gamma}_x q_x(0;p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p) \theta q_x(\theta;p) \Big).$$

To prove the point (ii), again by Lemma 3.8, we have that $\tau_n^\star(p) \xrightarrow{p} -\ln z^\star(p)$, so $e^{-\tau_n^\star(p)} \xrightarrow{p} z^\star(p)$. By a similar argument and applying Theorem 3.7, we conclude that

$$\Gamma_n(\tau_n^\star(p)) \xrightarrow{p} f_\Gamma(z^\star(p);p).$$

□

3.5.4 Proof of Theorem 3.11

By Theorem 3.9 and Corollary 3.10, for all $p \in [p_{\min}, p_0]$,

$$\kappa_n(p) \xrightarrow{p} g\big(f_\Gamma(z^\star(p);p)\big).$$

Let us define

$$\Phi_n(p) := p - g(\Gamma_n(\tau_n^\star;p)/n),$$

so that for a fixed $p_1 > \bar{p}$, we have

$$\Phi_n(p_1) = p_1 - g\big(f_\Gamma(z^\star(p_1); p_1)\big) - o_p(1).$$

From Definition (3.13) for $\bar{p}$, it follows that, for n large enough, $\mathbb{P}(p_n^\star > p_1) \to 0$.

Moreover, since $\bar{p}$ is a stable solution, when $\bar{p} = p_{\min}$, then by taking p_1 arbitrarily close to $p_{\min}$, we have that $p_n^\star \xrightarrow{p} \bar{p}$. When $\bar{p} \in (0,1]$, there exists $\epsilon > 0$ such that $p < g\big(f_\Gamma(z^\star(p); p)\big)$ for all $p \in (\bar{p} - \epsilon, \bar{p})$. Similarly, for any $\bar{p} - \epsilon < p_2 < \bar{p}$, we have $\Phi_n(p_2) < 0$ with high probability, i.e., as $n \to \infty$, $\mathbb{P}(p_n^\star < p_2) \to 0$. Then by taking p_1 and p_2 arbitrarily close to $\bar{p}$, we conclude that $p_n^\star \xrightarrow{p} \bar{p}$.

As seen above, when $z^\star(\bar{p}) = 0$, then $q_x(\infty; \bar{p}) = 0$. It follows that

$$g(f_\Gamma(z^\star(\bar{p}); \bar{p})) = g\Big(\sum_{x\in\mathcal{X}} \mu_x\big(\bar{\gamma}_x q_x(0; \bar{p}) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(\bar{p})\theta q_x(\theta; \bar{p})\big)\Big).$$

Moreover, as shown by Lemma 3.25 in Section 3.5.9 below, the function ϕ is locally continuous at $\bar{p}$. It follows that $\bar{p}$ is the largest solution of the fixed point equation

$$p = g\Big(\sum_{x\in\mathcal{X}} \mu_x\big(\bar{\gamma}_x q_x(0; p) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\theta q_x(\theta; p)\big)\Big).$$

This completes the proof of Theorem 3.11. □

3.5.5 Discussion on Lemma 3.6 and Lemma 3.13

We will discuss here how to extend the theorems presented in Chapter 2 to account for the possibility of an institution never defaulting, that is, an institution with an ∞ threshold.

We only consider the proof of limit theorem for $W_n(t; p)$ (Equation (3.9) and (3.14) for $i = W$); the proof of generalizations for $S_n(t; p)$ and $D_n(t; p)$ are similar. We denote by $L_n(t; p)$ and $H_n^-(t; p)$ the number of alive (not removed) outgoing half-edges at time t and the number of healthy (coming from solvent institutions) outgoing half-edges at time t respectively. From the definition of white outgoing half-edges process $W_n(t; p)$, it is clear that $W_n(t; p) = L_n(t; p) - H_n^-(t; p)$. Further,

$$H_n^-(t; p) = \sum_{x\in\mathcal{X}} d_x^- \big(\sum_{\theta=1} V_{x,\theta,d_x^+-\theta+1}^{(n)}(t; p) + N_{x,\infty}^{(n)}(p)\big).$$

We further denote by $\widetilde{W}_n(t; p)$ and $\widetilde{f}_W$ the outgoing half-edges process and corresponding limiting function as in Chapter 2. It is shown in Chapter 2 that

$$\widetilde{W}_n(t; p) = L_n(t; p) - \sum_{x\in\mathcal{X}} d_x^- \sum_{\theta=1} V_{x,\theta,d_x^+-\theta+1}^{(n)}(t; p),$$

and,

$$\widetilde{f}_W(z;p) = \lambda z - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=1}^{d_x^+} q_x(\theta;p)\beta(d_x^+, z, d_x^+ - \theta + 1).$$

We thus have

$$W_n(t;p) = \widetilde{W}_n(t;p) - \sum_{x\in\mathcal{X}} d_x^- N_{x,\infty}^{(n)} \quad \text{and} \quad f_W(z;p) = \widetilde{f}_W(z;p) - \sum_{x\in\mathcal{X}} d_x^- \mu_x q_x(\infty).$$

Further, as shown in Section 3.5.2,

$$\sum_{x\in\mathcal{X}} d_x^- N_{x,\infty}^{(n)}/n \xrightarrow{p} \sum_{x\in\mathcal{X}} d_x^- \mu_x q_x(\infty).$$

Together with Theorem 2.1 in Chapter 2, which gives

$$\sup_{t\leqslant\tau_n} \Big|\frac{\widetilde{W}_n(t;p)}{n} - \widetilde{f}_W(e^{-t};p)\Big| \xrightarrow{p} 0,$$

we obtain

$$\sup_{t\leqslant\tau_n} \Big|\frac{W_n(t;p)}{n} - f_W(e^{-t};p)\Big| \xrightarrow{p} 0,$$

which shows how to generalize the limit result of $W_n(t;p)$ in Lemma 3.6.

We next show how to generalize the asymptotic normality of $W_n(t;p)$, as in Lemma 3.13. We have

$$\begin{aligned} n^{-1/2}\Big(W_n(t\wedge\tau_n;p) - n\widehat{f}_W^{(n)}(t\wedge\tau_n;p)\Big) =& n^{-1/2}\Big(\widetilde{W}_n(t\wedge\tau_n;p) - n\widehat{\widetilde{f}}_W^{(n)}(t\wedge\tau_n;p)\Big) \\ &- \sum_{x\in\mathcal{X}} d_x^- n^{-1/2}\big(N_{x,\infty}^{(n)} - n\mu_x^{(n)} q_x^{(n)}(\infty)\big). \end{aligned}$$

By Lemma 3.23 and following similar arguments as in the proof of Theorem 3.14, one can show that the second term of the r.h.s. of the above formula is asymptotically Gaussian. The first term is also asymptotically Gaussian as shown in Theorem 2.6 in Chapter 2. Moreover, they are jointly asymptotically Gaussian. It remains to calculate the form of the variance function for $\sigma_W(e^{-t};p)$ of the limit white outgoing defaulted half-edges process. To do this, we write the limit process as

$$\mathcal{Z}_W(t;p) = \mathcal{Z}_L(t;p) - \sum_{x\in\mathcal{X}} d_x^- \Big(\sum_{\theta=1}^{d_x^+} \mathcal{Z}^*_{x,\theta,d_x^+-\theta+1}(t;p) + \mathcal{Y}^*_{x,\infty}\Big),$$

where $\mathcal{Z}_L(t;p)$ is the limit process for $n^{-1/2}(L_n(t;p) - n\lambda^{(n)}e^{-t})$. Moreover, as shown in Chapter 2, $\mathcal{Z}_L(t;p)$ is asymptotically Gaussian jointly with $\mathcal{Z}^*_{x,\theta,s}$ for all possible (x,θ,s) and jointly with $\mathcal{Y}^*_{x,\theta}$ for all (x,θ). Further, the covariances w.r.t. $\mathcal{Z}_L$ are also given in Chapter 2 by

$$\sigma_L(y) := \mathrm{Var}(\mathcal{Z}_L(-\ln y)) = \lambda(y - y^2)/2, \tag{3.19}$$

and

$$\begin{aligned}\sigma^L_{x,\theta,s}(y;p) :&= \mathrm{Cov}\big(\mathcal{Z}_L(-\ln y), \mathcal{Z}^*_{x,\theta,s}(-\ln y)\big)\\ &= y^{s+1}\sum_{j=s}^{d_x^+}\binom{j-1}{s-1}\int_y^1 (v-y)^{j-s}v^{-(j+1)}d\varphi_{x,\theta,j}(v;p).\end{aligned} \tag{3.20}$$

Notice that $\mathcal{Z}_L$ is independent of $\mathcal{Y}^*_{x,\theta}$ for all (x,θ). Then combining with the covariances given in Lemma 3.23, we conclude that

$$\begin{aligned}\sigma_W(y;p) = &\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\Big[(d_x^-)^2\widetilde{\sigma}_{x,\theta,\pi_x(\theta;p),\pi_x(\theta;p)}(y;p) - 2d_x^-\sigma^L_{x,\theta,\pi_x(\theta;p)}(y;p)\Big] + \sigma_L(y)\\ &+ \sum_{x\in\mathcal{X}}(d_x^-)^2\sum_{\theta_1=1}^{d_x^+}\sum_{\theta_2=1}^{d_x^+}\widehat{\sigma}_{x,\theta_1,\theta_2,\pi_x(\theta_1),\pi_x(\theta_2)}(y;p) + \sum_{x\in\mathcal{X}}(d_x^-)^2\psi_{x,\infty,\infty}(p)\\ &+ 2\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}(d_x^-)^2\beta(d_x^+, y, \pi_x(\theta;p))\psi_{x,\theta,\infty}(p),\end{aligned} \tag{3.21}$$

where $\pi_x(\theta;p) := d_x^+ - \theta + 1$, $\sigma_L(y)$ and $\sigma^L_{x,\theta,s}(y;p)$ are given by (3.19) and (3.20) respectively. Moreover, $\psi_{x,\theta,\infty}$, $\widetilde{\sigma}_{x,\theta,s_1,s_2}(y;p)$ and $\widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(y;p)$ are defined in Lemma 3.23.

3.5.6 Proof of Theorem 3.14

Recall that $U^{(n)}_{x,\theta,s}(t)$ denotes the number of institutions with type $x \in \mathcal{X}$, threshold θ and s alive incoming half-edges at time t. Further, we let $V^{(n)}_{x,\theta,s}(t)$ denote the number of institutions with type $x \in \mathcal{X}$, threshold θ and at least s alive incoming half-edges at time t, so that $V^{(n)}_{x,\theta,s}(t) = \sum_{\ell\geqslant s} U^{(n)}_{x,\theta,\ell}(t)$. We next define

$$V^{*(n)}_{x,\theta,s}(t;p) := n^{-1/2}\big(V^{(n)}_{x,\theta,s}(t;p) - n\mu^{(n)}_x q^{(n)}_x(\theta;p)\beta(d_x^+, e^{-t}, s)\big),$$

and

$$N^{*(n)}_{x,\theta}(p) := n^{-1/2}\big(N^{(n)}_{x,\theta}(p)(p) - n\mu^{(n)}_x q^{(n)}_x(\theta;p)\big).$$

We need the following result from Chapter 2, which shows the joint convergence of $N^{*(n)}_{x_1,\theta_1}$ and $V^{*(n)}_{x_2,\theta_2,s}$ for all possible (x_1,θ_1) and (x_2,θ_2,s). We recall that in the chapter we allow the threshold to be $\theta = \infty$ (see Section 3.5.5) and the results depend on p. But the lemma stays valid fo any fixed $p \in [p_{\min}, p_0]$.

Lemma 3.23. *Let $\tau_n \leqslant \tau_n^\star$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Under Assumption 3.5b and for any fixed $p \in [p_{\min}, p_0]$, we have that for all couple $x \in \mathcal{X}$ and $\theta \in \{1,\dots,d_x^+\}\cup\{\infty\}$, jointly as $n\to\infty$,*

$$N^{*(n)}_{x,\theta}(p) \xrightarrow{d} \mathcal{Y}^*_{x,\theta}(p),$$

where all $\mathcal{Y}^*_{x,\theta}(p)$ are Gaussian random variables with mean 0 and covariance

$$\operatorname{Cov}(\mathcal{Y}^*_{x_1,\theta_1}(p), \mathcal{Y}^*_{x_2,\theta_2}(p)) = \psi_{x_1,\theta_1,\theta_2}(p)\mathbb{1}\{x_1 = x_2\},$$

with

$$\psi_{x,\theta,\theta}(p) := \mu_x q_x(\theta;p)(1 - q_x(\theta;p)), \quad \psi_{x,\theta_1,\theta_2}(p) := -\mu_x q_x(\theta_1;p) q_x(\theta_2;p) \quad \textit{for all} \quad \theta_1 \neq \theta_2.$$

Further, for all triple (x,θ,s), jointly in $\mathcal{D}[0,\infty)$ and as $n \to \infty$,

$$V^{*(n)}_{x,\theta,s}(t \wedge \tau_n; p) \xrightarrow{d} \mathcal{Z}^*_{x,\theta,s}(t \wedge t_0; p),$$

where all $\mathcal{Z}^*_{x,\theta,s}(t;p)$ are continuous Gaussian processes with mean 0 and covariances

$$\begin{aligned}
\operatorname{Cov}\big(\mathcal{Z}^*_{x_1,\theta_1,s_1}(t;p), \mathcal{Z}^*_{x_2,\theta_2,s_2}(t;p)\big) &= 0, \quad \textit{for all} \quad x_1 \neq x_2,\\
\operatorname{Cov}\big(\mathcal{Z}^*_{x,\theta_1,s_1}(t;p), \mathcal{Z}^*_{x,\theta_2,s_2}(t;p)\big) &= \widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(e^{-t};p), \quad \textit{for all} \quad \theta_1 \neq \theta_2,\\
\operatorname{Cov}\big(\mathcal{Z}^*_{x,\theta,s_1}(t;p), \mathcal{Z}^*_{x,\theta,s_2}(t;p)\big) &= \widehat{\sigma}_{x,\theta,\theta,s_1,s_2}(e^{-t};p) + \widetilde{\sigma}_{x,\theta,s_1,s_2}(e^{-t};p),
\end{aligned}$$

with

$$\widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(e^{-t};p) := \beta(d^+_x, e^{-t}, s_1)\beta(d^+_x, e^{-t}, s_2)\psi_{x,\theta_1,\theta_2}(p),$$

$\widetilde{\sigma}_{x,\theta,s_1,s_2} = \widetilde{\sigma}_{x,\theta,s_2,s_1}$ *and*

$$\widetilde{\sigma}_{x,\theta,s,s+k}(y;p) := \frac{1}{2} y^{2s+k} \sum_{j=s+k}^{d^+_x} \binom{j-1}{s-1}\binom{j-1}{s+k-1} \int_y^1 (v-y)^{2j-2s-k} v^{-2j} d\varphi_{x,\theta,j}(v;p),$$

where $\varphi_{x,\theta,j}(y;p) := \mu_x q_x(\theta;p)\beta(d^+_x, y, j)$.

Moreover, the covariance between $\mathcal{Z}^*_{x_1,\theta_1,s}(t;p)$ and $\mathcal{Y}^*_{x_2,\theta_2}(p)$ is given by

$$\operatorname{Cov}\big(\mathcal{Z}^*_{x_1,\theta_1,s}(t;p), \mathcal{Y}^*_{x_2,\theta_2}(p)\big) = \beta(d^+_{x_1}, e^{-t}, s)\psi_{x_1,\theta_1,\theta_2}(p)\mathbb{1}\{x_1 = x_2\}.$$

By using the above lemma, we first show the following result regarding the asymptotic normality for $I_{x,\theta}(t;p)$, the total number of liquidations for institutions with type $x \in \mathcal{X}$ and threshold θ up to time t and under price p .

Lemma 3.24. *Let $\tau_n \leqslant \tau^\star_n$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Under Assumption 3.5b and for any fixed $p \in [p_{\min}, p_0]$, for all $x \in \mathcal{X}$, $\theta \in \{1,\ldots,d^+_x\} \cup \{\infty\}$, we have the following joint convergence in $\mathcal{D}[0,\infty)$ as $n \to \infty$,*

$$n^{-1/2}(I^{(n)}_{x,\theta}(t \wedge \tau_n; p) - n\widehat{f}^{(n)}_{x,\theta}(t \wedge \tau_n; p)) \xrightarrow{d} \mathcal{Z}_{I_{x,\theta}}(t \wedge t_0; p), \tag{3.22}$$

where all $\mathcal{Z}_{I_{x,\theta}}(t;p)$ are Gaussian processes with mean 0 and covariances

$$\operatorname{Cov}(\mathcal{Z}_{I_{x_1,\theta_1}}(t;p), \mathcal{Z}_{I_{x_2,\theta_2}}(t;p)) = \sigma^I_{x_1,\theta_1,\theta_2}(e^{-t};p)\mathbb{1}\{x_1 = x_2\},$$

where the form of $\sigma^I_{x,\theta_1,\theta_2}(y;p)$ is given by (3.28)-(3.31).

For the sake of readability, we postpone the proof of lemma to the end of this section.

We next consider the total liquidations, given by

$$Y_{x,\theta}^{(n)}(t;p) := \sum_{i=1}^{I_{x,\theta}^{(n)}(t;p)} L_{x,\theta}^{(i)}(p) \quad \text{and} \quad Y_{x,\infty}^{(n)}(t;p) := \sum_{i=1}^{I_{x,\infty}^{(n)}(t;p)} L_{x,\infty}^{(i)}(p),$$

where $\{L_{x,\theta}^{(i)}(p)\}_{i=1}^{\infty}$ are i.i.d. positive bounded random variables with expectation $\bar{\ell}_{x,\theta}(p)$ and variance $\varsigma_{x,\theta}^2(p)$ for $p \in [p_{\min}, p_0], x \in \mathcal{X}$ and $\theta \in \{1, \ldots, d_x^+\} \cup \{\infty\}$.

Note that conditioned on $I_{x_1,\theta_1}^{(n)}$ and $I_{x_2,\theta_2}^{(n)}$, the processes $Y_{x_1,\theta_1}^{(n)}(t;p)$ and $Y_{x_2,\theta_2}^{(n)}(t;p)$ are independent for $(x_1, \theta_1) \neq (x_2, \theta_2)$. In particular, from Lemma 3.24, for $x_1 \neq x_2$ we have that

$$\mathrm{Cov}\big(Y_{x_1,\theta_1}^{(n)}(t;p), Y_{x_2,\theta_2}^{(n)}(t;p)\big) = 0.$$

Consider now the decomposition

$$Y_{x,\theta}^{(n)}(t;p) - \widehat{f}_{x,\theta}^{(n)}(t;p) = \big(Y_{x,\theta}^{(n)}(t;p) - \bar{\ell}_{x,\theta}(p) I_{x,\theta}^{(n)}(t;p)\big) + \big(\bar{\ell}_{x,\theta}(p) I_{x,\theta}^{(n)}(t;p) - \widehat{f}_{x,\theta}^{(n)}(t;p)\big),$$

which implies that

$$\mathrm{Cov}\big(Y_{x,\theta_1}^{(n)}(t;p), Y_{x,\theta_2}^{(n)}(t;p)\big) = \bar{\ell}_{x,\theta_1}(p)\bar{\ell}_{x,\theta_2}(p)\mathrm{Cov}\big(I_{x,\theta_1}^{(n)}(t;p), I_{x,\theta_2}^{(n)}(t;p)\big),$$

and the same holds for their limit processes.

We now proceed to the proof of Theorem 3.14. The proof is based on a central limit theorem for processes which can be written as $Y_n(t) := \sum_{i=1}^{\lfloor X_n(t) \rfloor} G_i$, where $X_n(t)$ is a non-decreasing stochastic process satisfying $X_n(t) = O(n)$ for all $t > 0$ and $\{G_i\}_{i \geqslant 1}$ are i.i.d. positive bounded random variables. This is provided in Section 3.5.1.

Notice that the processes $I_{x,\theta}^{(n)}(t \wedge \tau_n^\star; p)$ for all $x \in \mathcal{X}, p \in [p_{\min}, p_0]$ and $\theta \in \{1, \ldots, d_x^+\} \cup \{\infty\}$ satisfy the conditions for $X_n(t)$ in Proposition 3.21. Indeed, $\widehat{f}_{x,\theta}^{(n)}(t;p) \to \widehat{f}_{x,\theta}(t;p)$ uniformly on $[0, \infty)$. Combining with the continuity of $\widehat{f}_{x,\theta}(t;p)$, it follows that $\widehat{f}_{x,\theta}^{(n)}(t \wedge \tau_n; p) \xrightarrow{p} \widehat{f}_{x,\theta}(t \wedge t_0)$, as $n \to \infty$. Using the Skorokhod coupling theorem [156, Theorem 3.30], we can assume that $\tau_n \to t_0$ a.s. in a new common probability space. It follows that a.s. $\widehat{f}_{x,\theta}^{(n)}(t \wedge \tau_n; p) \to \widehat{f}_{x,\theta}(t \wedge t_0)$. Thus, by Lemma 3.24, we have that for each ω outside a probability null set, for all $x \in \mathcal{X}$ and $\theta \in \{1, \ldots, d_x^+\} \cup \{\infty\}$, the process $I_{x,\theta}^{(n)}(t \wedge \tau_n; p)$ satisfies the conditions for $X_n(t)$ in Proposition 3.21, with $f_n(t) = \widehat{f}_{x,\theta}^{(n)}(t \wedge \tau_n(\omega))$ for different ω, but common $f(t) = \widehat{f}_{x,\theta}(t \wedge t_0)$ and $\mathcal{V} = \mathcal{Z}_{x,\theta}(t \wedge t_0)$. This leads to the same limit distribution up to a probability null set.

Let

$$\Delta_{x,\theta}^{(n)}(t;p) := n^{-1/2}\big(Y_{x,\theta}^{(n)}(t;p) - n\bar{\ell}_{x,\theta}\widehat{f}_{x,\theta}^{(n)}(t;p)\big).$$

By Proposition 3.21, we have that for all $x \in \mathcal{X}$, $\theta \in \{1, \ldots, d_x^+\} \cup \{\infty\}$ and a fixed $t > 0$, as $n \to \infty$, the following convergence holds

$$\Delta_{x,\theta}^{(n)}(t \wedge \tau_n; p) \xrightarrow{d} \mathcal{Z}_{x,\theta}(t \wedge t_0; p),$$

where $\mathcal{Z}_{x,\theta}(t;p)$ is Gaussian with mean 0 and variance

$$\Psi_{x,\theta}(t;p) := \widehat{f}_{x,\theta}(t;p)\varsigma^2_{x,\theta}(p) + \overline{\ell}^2_{x,\theta}(p)\sigma^I_{x,\theta,\theta}(e^{-t};p). \tag{3.23}$$

From the above arguments, the covariances between two different classes $x_1 \neq x_2$ are 0 and for $\theta_1 \neq \theta_2$, we have

$$\mathrm{Cov}(\mathcal{Z}_{x,\theta_1}(t;p), \mathcal{Z}_{x,\theta_2}(t;p)) = \bar{\ell}_{x,\theta_1}(p)\bar{\ell}_{x,\theta_2}(p)\sigma^I_{x,\theta_1,\theta_2}(e^{-t};p).$$

We next consider the convergence of the following infinite sum

$$\sum_{x\in\mathcal{X}}\sum_{\theta=1}^{d_x^+}\Delta^{(n)}_{x,\theta}(t\wedge\tau_n;p).$$

Recall that $\mathcal{X}_s^+$ denotes the collection of all classes $x \in \mathcal{X}$ with the in-degree $d_x^+ \geqslant s$. Recall also that all random variables $L^{(i)}_{x,\theta}(p)$ are assumed to be bounded. Then there exists some constant C such that $L^{(i)}_{x,\theta}(p) < C$, for all $x \in \mathcal{X}, p \in [p_{\min}, p_0], \theta \in \{1,\dots,d_x^+\} \cup \{\infty\}$ and $i \in \mathbb{N}$. Thus we have for any fixed $T > 0$,

$$\begin{aligned}\mathbb{E}\big[\sup_{t\leqslant T}\big|\sum_{x\in\mathcal{X}_s^+}\sum_{\theta=1}^{d_x^+}\Delta^{(n)}_{x,\theta}(t;p)\big|\big] &\leqslant C\mathbb{E}\big[\sup_{t\leqslant T}\big|\sum_{x\in\mathcal{X}_s^+}\sum_{\theta=1}^{d_x^+}n^{-1/2}(I^{(n)}_{x,\theta}(t;p) - n\widehat{f}^{(n)}_{x,\theta}(t;p))\big|\big]\\ &+ \mathbb{E}\big[\sup_{t\leqslant T}\big|\sum_{x\in\mathcal{X}_s^+}\sum_{\theta=1}^{d_x^+}n^{-1/2}(Y^{(n)}_{x,\theta}(t;p) - \bar{\ell}_{x,\theta}I^{(n)}_{x,\theta}(t;p))\big|\big].\end{aligned} \tag{3.24}$$

We first show that the first term of the r.h.s. inequality converges to 0 as $s \to \infty$ for n large enough. Indeed, results in Chapter 2 implies that when n is large enough, for any $T > 0$, as $\ell \to \infty$,

$$\mathbb{E}\big[\sup_{t\leqslant T}\big|\sum_{x\in\mathcal{X}_\ell^+}\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+}V^{*(n)}_{x,\theta,s}(t\wedge\tau_n;p)\big|\big] \to 0.$$

Moreover, as shown in the proof of Lemma 3.24,

$$\sum_{x\in\mathcal{X}_s^+}\sum_{\theta=1}^{d_x^+}n^{-1/2}(I^{(n)}_{x,\theta}(t;p) - n\widehat{f}^{(n)}_{x,\theta}(t;p)) = \sum_{x\in\mathcal{X}_\ell^+}\sum_{\theta=1}^{d_x^+}N^{*(n)}_{x,\theta} - \sum_{x\in\mathcal{X}_\ell^+}\sum_{\theta=1}^{d_x^+}\sum_{s=d_x^+-\theta+1}^{d_x^+}V^{*(n)}_{x,\theta,s}(t\wedge\tau_n;p),$$

and by Cauchy-Schwarz inequality,

$$\mathbb{E}\big|\sum_{x\in\mathcal{X}_\ell^+}\sum_{\theta=1}^{d_x^+}N^{*(n)}_{x,\theta}\big| \leqslant \sum_{x\in\mathcal{X}_\ell^+}q^{(n)}_x(\infty)(1-q^{(n)}_x(\infty)),$$

which goes to 0 as $\ell \to \infty$ uniformly in n. We conclude that the first term of the r.h.s. of (3.24) converges to 0. For the second term first note that each term of the sum inside the expectation is a martingale. Then, by Doob inequality, we can control its L^2-norm by $4C^2 \sum_{x\in\mathcal{X}_s^+} d_x^+ \mu_x^{(n)}$. Hence, using Assumption 3.5b, the L^2-bound converges to 0 as $s \to \infty$ for n large enough, and the second term converges to 0 as desired. We can then take the limit under the infinite sum, by using e.g., [64, Theorem 4.2]. It follows that

$$\sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \Delta_{x,\theta}^{(n)}(t \wedge \tau_n; p) \xrightarrow{d} \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \mathcal{Z}_{x,\theta}(t \wedge t_0).$$

For the second and third term of $n^{-1/2}(\Gamma_n(t \wedge \tau_n; p) - n\widehat{f}_\Gamma^{(n)}(t \wedge \tau_n; p))$, by using similar arguments as above, we obtain

$$\sum_{x\in\mathcal{X}} n^{-1/2}\big(Y_{x,\infty}^{(n)}(t;p) - n\bar{\ell}_{x,\infty}\widehat{f}_{x,\infty}^{(n)}(t \wedge \tau_n; p)\big) \xrightarrow{d} \sum_{x\in\mathcal{X}} \mathcal{Z}_{x,\infty}(t \wedge t_0),$$

and,

$$\sum_{x\in\mathcal{X}} \bar{\gamma}_x N_{x,0}^{*(n)} \xrightarrow{d} \sum_{x\in\mathcal{X}} \bar{\gamma}_x \mathcal{Y}_{x,0}^*.$$

Hence we have

$$\mathcal{Z}_\Gamma(t \wedge t_0) := \sum_{x\in\mathcal{X}} \sum_{\theta=1}^{d_x^+} \mathcal{Z}_{x,\theta}(t \wedge t_0) + \sum_{x\in\mathcal{X}} \mathcal{Z}_{x,\infty}(t \wedge t_0) + \sum_{x\in\mathcal{X}} \bar{\gamma}_x \mathcal{Y}_{x,0}^*,$$

which is a centered Gaussian random variable with mean 0. By Lemma 3.23, Lemma 3.24 and above arguments, the variance is given by

$$\begin{aligned}
\Psi(t;p) = &\sum_{x\in\mathcal{X}} \big(2\sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\bar{\ell}_{x,\infty}(p)\sigma_{x,\theta,\infty}^I(e^{-t};p) - 2\bar{\gamma}_x \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p)\psi_{x,\theta,0}(p) \sum_{s=d_x^+-\theta+1}^{d_x^+} \beta(d_x^+, e^{-t}, s)\big) \\
&+ \sum_{x\in\mathcal{X}} \sum_{\theta_1,\theta_2=1}^{d_x^+} \bar{\ell}_{x,\theta_1}(p)\bar{\ell}_{x,\theta_2}(p)\sigma_{x,\theta_1,\theta_2}^I(e^{-t};p) - \sum_{x\in\mathcal{X}} 2\bar{\gamma}_x\bar{\ell}_{x,\infty}(p)\psi_{x,\infty,0}(p) \sum_{s=1}^{d_x^+} \beta(d_x^+, e^{-t}, s) \\
&+ \sum_{x\in\mathcal{X}} \big(\sum_{\theta=1}^{d_x^+} \widehat{f}_{x,\theta}(t;p)\varsigma_{x,\theta}^2(p) + \Psi_{x,\infty}(t;p) + \bar{\gamma}_x^2\psi_{x,0,0}(p)\big) \\
&+ 2\sum_{x\in\mathcal{X}} \big(\sum_{\theta=1}^{d_x^+} \theta\bar{\ell}_{x,\theta}(p)\bar{\gamma}_x\psi_{x,\theta,0}(p) + d_x^+\bar{\ell}_{x,\infty}(p)\bar{\gamma}_x\psi_{x,\infty,0}(p)\big)
\end{aligned} \tag{3.25}$$

where $\psi_{x,\theta_1,\theta_2}$ is defined in Lemma 3.23, $\sigma_{x,\theta_1,\theta_2}^I$ is given by (3.28)-(3.31) and $\Psi_{x,\theta}$ is defined by (3.23).

We are left to prove Lemma 3.24.

Proof of Lemma 3.24. Recall that $V^{(n)}_{x,\theta,s}$ denotes the number of institutions with type x, threshold θ and with at least s incoming half-edges at time t. We have

$$\begin{aligned} I^{(n)}_{x,\theta}(t;p) =& \theta N^{(n)}_{x,\theta}(p) - \sum_{s=d^+_x-\theta+1}^{d^+_x} (s - d^+_x + \theta) U^{(n)}_{x,\theta,s}(t;p) \\ =& \theta N^{(n)}_{x,\theta}(p) - \sum_{s=d^+_x-\theta+1}^{d^+_x} V^{(n)}_{x,\theta,s}(t;p). \end{aligned}$$

It follows that

$$n^{-1/2}(I^{(n)}_{x,\theta}(t \wedge \tau_n;p) - n\widehat{f}^{(n)}_{x,\theta}(t \wedge \tau_n)) = \theta N^{*(n)}_{x,\theta} - \sum_{s=d^+_x-\theta+1}^{d^+_x} V^{*(n)}_{x,\theta,s}(t \wedge \tau_n;p), \tag{3.26}$$

and,

$$n^{-1/2}(I^{(n)}_{x,\infty}(t \wedge \tau_n;p) - n\widehat{f}^{(n)}_{x,\infty}(t \wedge \tau_n)) = d^+_x N^{*(n)}_{x,\infty} - \sum_{s=1}^{d^+_x} V^{*(n)}_{x,\theta,s}(t \wedge \tau_n;p). \tag{3.27}$$

By Lemma 3.23, we have the joint convergence of $N^{*(n)}_{x_1,\theta_1}$ and $V^{*(n)}_{x_2,\theta_2,s}$ for all possible (x_1,θ_1) and (x_2,θ_2,s). We therefore have for $\theta \in \{1,\dots,d^+_x\}$,

$$\mathcal{Z}_{I_{x,\theta}}(t;p) := \theta \mathcal{Y}^*_{x,\theta} - \sum_{s=d^+_x-\theta+1}^{d^+_x} \mathcal{Z}^*_{x,\theta,s}(t;p),$$

and for the threshold $\theta = \infty$,

$$\mathcal{Z}_{I_{x,\infty}}(t;p) := d^+_x \mathcal{Y}^*_{x,\infty} - \sum_{s=1}^{d^+_x} \mathcal{Z}^*_{x,\theta,s}(t;p).$$

By using the covariance formulas in Lemma 3.23 and some basic calculations, we obtain the following formulas for the covariance $\sigma^I_{x,\theta_1,\theta_2}(e^{-t};p)$.

- For $\theta_1 = \theta_2 = \theta \in \{1,\dots,d^+_x\}$:

$$\begin{aligned} \sigma^I_{x,\theta,\theta}(y;p) =& \theta^2 \psi_{x,\theta,\theta}(p) + \sum_{s_1,s_2=d^+_x-\theta+1}^{d^+_x} (\widehat{\sigma}_{x,\theta,\theta,s_1,s_2}(y;p) + \widetilde{\sigma}_{x,\theta,s_1,s_2}(y;p)) \\ & - 2\theta\psi_{x,\theta,\theta}(p) \sum_{s=d^+_x-\theta+1}^{d^+_x} \beta(d^+_x, y, s), \end{aligned} \tag{3.28}$$

- For $\theta_1, \theta_2 \in \{1, \dots, d_x^+\}$ and $\theta_1 \neq \theta_2$:

$$
\begin{aligned}
\sigma^I_{x,\theta_1,\theta_2}(y;p) =& \theta_1\theta_2\psi_{x,\theta_1,\theta_2}(p) + \sum_{s_1=d_x^+-\theta_1+1}^{d_x^+} \sum_{s_2=d_x^+-\theta_2+1}^{d_x^+} \widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}(y;p) \\
&- \theta_1\psi_{x,\theta_1,\theta_2}(p) \sum_{s=d_x^+-\theta_2+1}^{d_x^+} \beta(d_x^+, y, s) - \theta_2\psi_{x,\theta_1,\theta_2}(p) \sum_{s=d_x^+-\theta_1+1}^{d_x^+} \beta(d_x^+, y, s),
\end{aligned}
\tag{3.29}
$$

- For $\theta_1 = \theta_2 = \infty$:

$$
\begin{aligned}
\sigma^I_{x,\infty,\infty}(y;p) =& (d_x^+)^2\psi_{x,\infty,\infty}(p) + \sum_{s_1,s_2=1}^{d_x^+} \left(\widehat{\sigma}_{x,\theta,\theta,s_1,s_2}(y;p) + \widetilde{\sigma}_{x,\theta,s_1,s_2}(y;p)\right) \\
&- 2d_x^+\psi_{x,\infty,\infty}(p) \sum_{s=1}^{d_x^+} \beta(d_x^+, y, s),
\end{aligned}
\tag{3.30}
$$

- For $\theta_1 = \infty$ and $\theta_2 = \theta \in \{1, \dots, d_x^+\}$:

$$
\begin{aligned}
\sigma^I_{x,\infty,\theta}(y;p) =& d_x^+\theta\psi_{x,\infty,\theta}(p) + \sum_{s_1=1}^{d_x^+} \sum_{s_2=d_x^+-\theta+1}^{d_x^+} \widehat{\sigma}_{x,\infty,\theta,s_1,s_2}(y;p) \\
&- d_x^+\psi_{x,\infty,\theta}(p) \sum_{s=d_x^+-\theta+1}^{d_x^+} \beta(d_x^+, y, s) - \theta\psi_{x,\infty,\theta}(p) \sum_{s=1}^{d_x^+} \beta(d_x^+, y, s),
\end{aligned}
\tag{3.31}
$$

where the forms of $\widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}$, $\widetilde{\sigma}_{x,\theta,s_1,s_2}$ and $\psi_{x,\theta_1,\theta_2}$ for all $\theta_1, \theta_2 \in \{1, \dots, d_x^+\} \cup \infty$ are provided in Lemma 3.23. This completes the proof of Lemma 3.24. $\square$

3.5.7 Proof of Theorem 3.16

Consider $z^\star(p) \in (0, 1]$ and $z^\star(p)$ is a stable solution, i.e., $\alpha(p) := f^1_W(z^\star(p); p) > 0$.

First note that the variance $\Psi(t; p)$ of $\mathcal{Z}_\Gamma(t; p)$ is continuous in t. Indeed, from the explicit forms of $\widetilde{\sigma}_{x,\theta,s_1,s_2}$ and $\widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}$ in Lemma 3.23, we have the following inequalities,

$$
|\widehat{\sigma}_{x,\theta_1,\theta_2,s_1,s_2}| \leqslant \mu_x
$$

and,

$$
|\widetilde{\sigma}_{x,\theta,s_1,s_2}| \leqslant \sum_{j=0}^{d_x^+} \int_y^1 \frac{y^2}{v^2} d\varphi_{x,\theta,j}(v;p) \leqslant 2d_x^+\mu_x q_x(\theta;p),
$$

for all $(x,\theta_1,\theta_2,s_1,s_2)$. Thus, we obtain that for all $x \in \mathcal{X}$, $\theta_1,\theta_2 \in \{1,\dots,d_x^+\} \cup \{\infty\}$,

$$\sigma^I_{x,\theta_1,\theta_2} \leqslant 4(d_x^+)^3\mu_x.$$

By the definition of $\Psi(t;p)$ as in (3.23), we have that for some constant C, the infinite tail sum of the first term in Ψ in (3.25) satisfies that

$$\sum_{x\in\mathcal{X}_\ell^+}\Big(\sum_{\theta=1}^{d_x^+}\Psi_{x,\theta}(t;p)+\Psi_{x,\infty}(t;p)+\bar{\gamma}_x^2\psi_{x,0,0}(p)\Big) \leqslant C\sum_{x\in\mathcal{X}_\ell^+}(d_x^+)^4\mu_x,$$

which goes to 0 as $\ell\to\infty$ by Assumption 3.5b. One can show by a similar argument that the other sum terms in Ψ have the same tail convergence property. Since each single term is continuous in t, again we can pass the continuity in the infinite sum. Moreover, since $\mathcal{Z}_\Gamma(t;p)$ is a centered Gaussian random variable, its distribution is determined by $\Psi(t;p)$. Thus for a sequence $\{t_n\}_n$ which converges to t, we have that as $n\to\infty$,

$$\mathcal{Z}_\Gamma(t_n;p) \xrightarrow{d} \mathcal{Z}_\Gamma(t;p). \tag{3.32}$$

Then we can use the Skorokhod representation theorem, which shows that one can change the probability space where all the random variables are well defined and all the convergence results of Theorem 3.14, Lemma 3.13, Lemma 3.8 ($\tau_n^\star\to t^\star$) and (3.32) hold almost surely. Taking $t=\tau_n^\star$ and $t_0=t^\star$, we obtain by Lemma 3.13 and by continuity of $\mathcal{Z}_W$ that

$$\begin{aligned}W_n(\tau_n^\star;p) &= n\widehat{f}_W^{(n)}(\tau_n^\star;p)+n^{1/2}\mathcal{Z}_W(\tau_n^\star\wedge t^\star;p)+o(n^{1/2})\\ &= n\widehat{f}_W^{(n)}(\tau_n^\star;p)+n^{1/2}\mathcal{Z}_W(t^\star;p)+o(n^{1/2}).\end{aligned}$$

Since $W_n(\tau_n^\star;p)=-1$, then

$$\widehat{f}_W^{(n)}(\tau_n^\star;p) = -n^{-1/2}\mathcal{Z}_W\big(t^\star;p\big)+o\big(n^{-1/2}\big).$$

Since, as $n\to\infty$, $\tau_n^\star\to t^\star$ and $t_n^\star\to t^\star$ hold a.s., there exists some ξ_n in the interval between $t_n^\star$ and $\tau_n^\star$ such that $\xi_n\to t^\star$. Further, as $n\to\infty$,

$$(\widehat{f}_W^{(n)})'(\xi_n;p)\to\widehat{f}_W'(t^\star;p) = -z^\star(p)\alpha(p).$$

It follows then by Mean-Value theorem that

$$\widehat{f}_W^{(n)}(\tau_n^\star;p)=\widehat{f}_W^{(n)}(\tau_n^\star;p)-\widehat{f}_W^{(n)}(t_n^\star;p)=(\widehat{f}_W^{(n)})'(\xi_n)(\tau_n^\star-t_n^\star)=(-z^\star(p)\alpha(p)+o(1))(\tau_n^\star-t_n^\star).$$

Hence we have

$$\tau_n^\star-t_n^\star=\Big(-\frac{1}{z^\star(p)\alpha(p)}+o(1)\Big)\widehat{f}_W^{(n)}(\tau_n^\star;p)=n^{-1/2}\frac{1}{z^\star(p)\alpha(p)}(Z_W(t^\star;p)+o(1)). \tag{3.33}$$

Moreover, it follows by Theorem 3.14 that

$$n^{-1/2}\Gamma_n(\tau_n^\star;p)=n^{1/2}\widehat{f}_\Gamma^{(n)}(\tau_n^\star;p)+\mathcal{Z}_\Gamma(\tau_n^\star\wedge t^\star;p)+o(1).$$

Since, as $n \to \infty$, $\tau_n^\star \wedge t^\star \to t^\star$ a.s., we obtain that a.s. $\mathcal{Z}_\Gamma(t_n) \to \mathcal{Z}_\Gamma(t;p)$. It then follows that, for some $\xi_n' \to t^\star$ as $n \to \infty$, that

$$\begin{aligned} n^{-1/2}\Gamma_n(\tau_n^\star;p) &= n^{1/2}\widehat{f}_\Gamma^{(n)}(\tau_n^\star;p) + \mathcal{Z}_\Gamma(t^\star;p) + o(1) \\ &= n^{1/2}\widehat{f}_\Gamma^{(n)}(\tau_n^\star;p) + n^{1/2}(\widehat{f}_\Gamma^{(n)})'(\xi_n';p)(\tau_n^\star - t_n^\star) + \mathcal{Z}_\Gamma(t^\star;p) + o(1). \end{aligned}$$

By plugging (3.33) into the above formula and doing some simplification, it follows that

$$n^{-1/2}\Gamma_n(\tau_n^\star;p) = n^{1/2} f_\Gamma^{(n)}(z_n^\star;p) - \frac{f_\Gamma'(z^\star;p)}{\alpha}\mathcal{Z}_W(t^\star;p) + \mathcal{Z}_\Gamma(t^\star;p) + o(1).$$

This completes the proof of Theorem 3.16. ☐

3.5.8 Proof of Theorem 3.18

Consider $z^\star(p) \in (0,1]$ and $z^\star(p)$ is a stable solution, i.e., $\alpha(p) := f_W^1(z^\star(p);p) > 0$. We have by Theorem 3.16 that $\Gamma_n(\tau_n^\star;p)$ is asymptotic normal

$$n^{1/2}(\Gamma_n(\tau_n^\star;p)/n - f_\Gamma^{(n)}(z_n^\star;p)) \xrightarrow{d} \mathcal{Z}_\Gamma(t^\star;p) - \alpha(p)^{-1} f_\Gamma'(z^\star;p)\mathcal{Z}_W(t^\star;p). \tag{3.34}$$

Since for any fixed $p \in [p_{\min}, p_0]$, $f_W^{(n)}(z;p)$ converges to $f_W(z;p)$ uniformly on $[0,1]$, we have that $z_n^\star(p) \to z^\star(p)$ as $n \to \infty$ in probability. Moreover, by continuity of f_Γ and $f_\Gamma^{(n)}$ for all n and the uniformly convergence of $f_\Gamma^{(n)}(\cdot;p)$ to $f_\Gamma(\cdot;p)$ for any fixed p, we can conclude that

$$f_\Gamma^{(n)}(z_n^\star;p) \to f_\Gamma(z^\star;p)$$

in probability for any $p \in [p_{\min}, p_0]$.

Since the inverse demand function g is in $\mathcal{C}^1$ by Assumption 3.1, we have

$$g'\big(f_\Gamma^{(n)}(z_n^\star;p)\big) \xrightarrow{p} g'\big(f_\Gamma(z^\star;p)\big).$$

By the mean-value theorem, there exists some ξ_n between $\Gamma_n(\tau_n^\star;p)/n$ and $f_\Gamma^{(n)}(z_n^\star;p)$ such that

$$g(\Gamma_n(\tau_n^\star;p)/n) - g\big(f_\Gamma^{(n)}(z_n^\star;p)\big) = g'(\xi_n)(\Gamma_n(p)/n - f_\Gamma^{(n)}(z_n^\star;p)). \tag{3.35}$$

Note that $\Gamma_n(\tau_n^\star;p)/n \xrightarrow{p} f_\Gamma(z^\star;p)$, thus we also have

$$g'(\xi_n) \xrightarrow{p} g'\big(f_\Gamma(z^\star;p)\big).$$

Multiplying both side of (3.35) by $n^{1/2}$ gives

$$n^{1/2}\big(h(\Gamma_n(\tau_n^\star;p)/n) - g(f_\Gamma^{(n)}(z_n^\star;p))\big) = n^{1/2} g'(\xi_n)(\Gamma_n(\tau_n^\star;p)/n - f_\Gamma^{(n)}(z_n^\star;p)).$$

By the asymptotic normality in (3.34) and Slutsky's theorem we obtain

$$n^{1/2}(\kappa_n(p) - g(f_\Gamma^{(n)}(z_n^\star;p))) \xrightarrow{d} g'(f_\Gamma(z^\star;p))\big(\mathcal{Z}_\Gamma(t^\star;p) - \alpha^{-1} f_\Gamma^1(z^\star;p)\mathcal{Z}_W(t^\star;p)\big).$$

This completes the proof of Theorem 3.18. ☐

3.5.9 Proof of Theorem 3.19

We first state a lemma which is used in the proof of Theorem 3.19. Let us define

$$\phi(p) := p - g \circ f_\Gamma(z^\star(p);p) \quad \text{and} \quad \phi_n(p) := p - g \circ f_\Gamma^{(n)}(z_n^\star(p);p).$$

Lemma 3.25. *Under Assumption 3.3 and Assumption 3.4, the following holds:*

(a) *For any fixed $p \in (p_{\min}, p_0)$, if $z^\star(p) = 0$ or $z^\star(p) \in (0,1)$ and $\alpha(p) > 0$, then there exists some small $\delta > 0$ and N large enough, such that $z^\star(\cdot)$ and all $z_n^\star(\cdot)$ for $n > N$ are continuous in the interval $(p-\delta, p+\delta)$;*

(b) *For $p \in \{p_{\min}, p_0\}$, with the same conditions as in (a), the continuities hold but on a semi-interval $[p_{\min}, p_{\min}+\delta)$ for $p = p_{\min}$ and $(p_0 - \delta, p_0]$ for $p = p_0$.*

(c) *If $\bar{p}$ is a stable fixed point solution, then under the same conditions as in (a), we have that, for N large enough, $\bar{p}$ and all $\{\bar{p}_n, n > N\}$ are continuity points of ϕ and ϕ_n, respectively. Moreover, as $n \to \infty$, $\bar{p}_n \to \bar{p}$.*

For the sake of readability, we postpone the proof of lemma to the end of this section and proceed with the proof of Theorem 3.19.

Consider now $z^\star(\bar{p}) \in (0,1]$ is a stable solution of $f_W(z;\bar{p}) = 0$, i.e., $\alpha(\bar{p}) := f_W^1(z^\star;\bar{p}) > 0$, and $\bar{p}$ is a stable solution of Equation (3.13).

By Lemma 3.25, we know that, for n large enough, $\bar{p}_n$ exists and converges to $\bar{p}$ as $n \to \infty$. Moreover, by Theorem 3.18, we have that as $n \to \infty$,

$$\Phi_n(p) - \phi_n(p) \xrightarrow{d} \mathcal{Z}_V(p).$$

$\mathcal{Z}_V(p)$ is a centered Gaussian random variable, and its distribution is determined uniquely by its variance. By the analysis in the proof of Theorem 3.16, the variance function of $\mathcal{Z}_\Gamma(t;p)$ is continuous in p. By similar arguments, the variance function of $\mathcal{Z}_W$ is also continuous in p. Then by Cauchy-Schwarz inequality, we can control the covariance between $\mathcal{Z}_\Gamma$ and $\mathcal{Z}_W$ by their variances. Thus the variance function of $\mathcal{Z}_V(p)$ is continuous in p. We therefore have that

$$\mathcal{Z}_V(p_n) \xrightarrow{d} \mathcal{Z}_V(p), \tag{3.36}$$

for any sequence $\{p_n\}_n$ which converges to p as $n \to \infty$.

The Skorokhod representation theorem shows that one can change the probability space where all the random variables are well defined and, all the convergence results of Theorem 3.18, the convergence in probability $p_n^\star \to \bar{p}$ and (3.36) hold a.s.. Then we can write

$$\begin{aligned}\Phi_n(p_n^\star) &= \phi_n(p_n^\star) + n^{-1/2}\mathcal{Z}_V(p_n^\star) + o(n^{-1/2}) \\ &= \phi_n(p_n^\star) + n^{-1/2}\mathcal{Z}_V(\bar{p}) + o(n^{-1/2}),\end{aligned} \tag{3.37}$$

where the second equality follows from $\mathcal{Z}_V(p_n^\star) \to \mathcal{Z}_V(\bar{p})$ a.s.. From $\Phi_n(p_n^\star) = 0$, we have

$$\phi_n(p_n^\star) = -n^{-1/2}\mathcal{Z}_V(\bar{p}) + o(n^{-1/2}). \tag{3.38}$$

Moreover, as $n \to \infty$, we have a.s. $p_n^\star \to \bar{p}$ and $\bar{p}_n \to \bar{p}$. Combining the continuity of $f_\Gamma^{(n)}$ and the local continuity of $z^\star(\cdot)$, we have that both $f_\Gamma^{(n)}(z_n^\star(p_n^\star); p_n^\star)$ and $f_\Gamma^{(n)}(z_n^\star(\bar{p}_n); \bar{p}_n)$ converge a.s. to $f_\Gamma(z^\star(\bar{p}), \bar{p})$. Thus, by the Mean-Value theorem, there exists some sequence $\{\xi_n\}$ with $\xi_n \to f_\Gamma(z^\star(\bar{p}); \bar{p})$ a.s. in the interval between $f_\Gamma^{(n)}(z_n^\star(p_n^\star); p_n^\star)$ and $f_\Gamma^{(n)}(z_n^\star(\bar{p}_n); \bar{p}_n)$ such that

$$g\big(f_\Gamma^{(n)}(z_n^\star(p_n^\star); p_n^\star)\big) - g\big(f_\Gamma^{(n)}(z_n^\star(\bar{p}_n); \bar{p}_n)\big) = g'(\xi_n)(f_\Gamma^{(n)}(z_n^\star(p_n^\star); p_n^\star) - f_\Gamma^{(n)}(z_n^\star(\bar{p}_n); \bar{p}_n)). \tag{3.39}$$

We next analyze $f_\Gamma^{(n)}(z_n^\star(p_n^\star); p_n^\star) - f_\Gamma^{(n)}(z_n^\star(\bar{p}_n); \bar{p}_n)$. By the Mean-Value theorem and Lemma 3.25, there exists a sequence $\{\xi_n^z\}$ and $\{\xi_n^p\}$ with $\xi_n^z \to z^\star(\bar{p})$ a.s and $\xi_n^p \to \bar{p}$ a.s. such that

$$f_\Gamma^{(n)}(z_n^\star(p_n^\star); p_n^\star) - f_\Gamma^{(n)}(z_n^\star(\bar{p}_n); \bar{p}_n) = f_\Gamma^{1,(n)}(\xi_n^z; p_n^\star)(z_n^\star(p_n^\star) - z_n^\star(\bar{p}_n)) + f_\Gamma^{2,(n)}(z_n^\star(\bar{p}_n); \xi_n^p)(p_n^\star - \bar{p}_n). \tag{3.40}$$

It remains to analyze $z_n^\star(p_n^\star) - z_n^\star(\bar{p}_n)$. Notice that, by definition, $f_W^{(n)}(z_n^\star(p); p) = 0$ for any $p \in [p_{\min}, p_0]$. By using again the Mean-Value theorem, we have the following relations

$$-f_W^{(n)}(z_n^\star(\bar{p}_n); p_n^\star) = f_W^{(n)}(z_n^\star(p_n^\star); p_n^\star) - f_W^{(n)}(z_n^\star(\bar{p}_n); p_n^\star) = f_W^{1,(n)}(\alpha_n^z; p_n^\star)(z_n^\star(p_n^\star) - z_n^\star(\bar{p}_n)),$$

and,

$$f_W^{(n)}(z_n^\star(\bar{p}_n); p_n^\star) = f_W^{(n)}(z_n^\star(\bar{p}_n); p_n^\star) - f_W^{(n)}(z_n^\star(\bar{p}_n; \bar{p}_n) = f_W^{2,(n)}(z_n^\star(\bar{p}_n); \alpha_n^p)(p_n^\star - \bar{p}_n),$$

where $\alpha_n^z \to z^\star(\bar{p})$ a.s. and $\alpha_n^p \to \bar{p}$ a.s. as $n \to \infty$. Then by above two equations we have

$$z_n^\star(p_n^\star) - z_n^\star(\bar{p}_n) = -(f_W^{1,(n)}(\alpha_n^z; p_n^\star))^{-1} f_W^{2,(n)}(z_n^\star(\bar{p}_n); \alpha_n^p)(p_n^\star - \bar{p}_n). \tag{3.41}$$

Now combining (3.39), (3.40) and (3.41) and using Remark 3.15, we obtain

$$\begin{aligned}
\phi_n(p_n^\star) &= \phi_n(p_n^\star) - \phi_n(\bar{p}_n) \\
&= p_n^\star - \bar{p}_n - (g'(f_\Gamma(z^\star(\bar{p}); \bar{p}) + o(1))[(f_\Gamma^1(z^\star(\bar{p}); \bar{p}) + o(1)) \\
&\qquad (-(f_W^1(z^\star(\bar{p}); \bar{p})^{-1} f_W^2(z^\star(\bar{p}); \bar{p}) + o(1)) + (f_\Gamma^2(z^\star(\bar{p}); \bar{p}) + o(1))](p_n^\star - \bar{p}_n) \\
&= (\rho + o(1))(p_n^\star - \bar{p}_n).
\end{aligned}$$

Using (3.38), we conclude

$$p_n^\star - \bar{p}_n = \Big(\frac{1}{\rho} + o(1)\Big)\phi_n(p_n^\star) = -n^{-1/2}\frac{1}{\rho}\mathcal{Z}_V(\bar{p}) + o(n^{-1/2}).$$

This completes the proof of Theorem 3.19. □

We are only left to prove Lemma 3.25.

Proof of Lemma 3.25. From the definition of the threshold distribution, $q_x^{(n)}(\theta;p)$ are (stochastically) non-decreasing on p for every (x,θ) and every n. Thus, for an increasing sequence p_n converging to some $p \in [p_{\min}, p_0]$, we can show that for any fixed $z \in [0,1]$, the sequence $\{f_W(z;p_n)\}_n$ is monotone and converges to $f_W(z;p)$. In addition, $f_W(z;p)$ and $f_W^{(n)}(z;p)$ are continuous in z for all n. It therefore follows by Dini's theorem that $\{f_W(\cdot;p_n)\}_n$ converges uniformly to $f_W(z;p)$ on $[0,1]$. Hence the largest root $z^\star(p_n)$ must also converge to $z^\star(p)$. The same argument for a decreasing sequence p_n gives the same uniform convergence. Thus $\{f_W(\cdot;p_n)\}_n$ converges uniformly to $f_W(\cdot;p)$ for any sequence converging to p.

If $z^\star(p) = 0$ or $z^\star(p) \in (0,1)$ and $\alpha(p) > 0$, then for some $\epsilon < \delta$ small enough, we have $f_W(z^\star(p) + \epsilon;p) > 0$ and $f_W(z^\star(p) - \epsilon;p) < 0$. Then for n large enough, it follows that $f_W(z^\star(p) + \epsilon;p_n) > 0$ and $f_W(z^\star(p) - \epsilon;p_n) < 0$. We therefore have $z^\star(p_n) \in (z^\star(p) - \epsilon, z^\star(p) + \epsilon)$, and ϵ can be arbitrarily small, thus $z^\star(p_n) \to z^\star(p)$ as $p_n \to p$. If $z^\star(p) = 0$, we have for some $\epsilon > 0$ that $f_W(z;p) > 0$ for all $z \geqslant \epsilon$. By the uniform convergence of p_n to p, for n large enough, we also have $f_W(z;p_n) > 0$ for $z \geqslant \epsilon$. Thus $z^\star(p_n) \in [0,\epsilon)$. Taking ϵ arbitrarily small, we conclude that $z^\star(p_n) \to z^\star(p)$ as $p_n \to p$. This continuity holds on a small interval $(p-\delta, p+\delta)$ for some δ small enough. A similar argument gives the same conclusion for the point (b).

It is also clear that for any fixed p, $f_W^{(n)}(z;p)$ converges to $f_W(z;p)$ point wisely on z. Since for any $z \in [0,1]$,

$$f_W^{(n)}(z;p) \leqslant \lambda^{(n)} + \sum_{x\in\mathcal{X}} \mu_x^{(n)} d_x^- \Big[\sum_{\theta=1}^{d_x^+} q_x^{(n)}(\theta;p) + q_x^{(n)}(\infty;p)\Big],$$

by Assumption 3.4 and applying dominated convergence theorem, we have further that $f_W^{(n)}(z;p)$ converges to $f_W(z;p)$ uniformly on z in $[0,1]$. The same argument applied to $f_\Gamma^{(n)}$ gives the uniform convergence of $f_\Gamma^{(n)}$ to f_Γ on z. By the uniform convergence of $f_W^{(n)}$ to f_W, it is obvious that we can choose $\epsilon < \delta$ such that the local continuity of $z^\star(\cdot)$ and of all $z_n^\star(\cdot)$ hold on $(\bar{p}-\epsilon, \bar{p}+\epsilon)$ for n large enough. This completes the proof of point (a) and (b).

We next proceed with the proof of point (c) of the lemma. We first prove the local continuity of ϕ on an interval where we assume that $z^\star(\cdot)$ is continuous in p. Recall that $\mathcal{X}_s^+$ is the collection of all classes $x \in \mathcal{X}$ with the in-degree $d_x^+ \geqslant s$. Since all $\bar{\ell}_{x,\theta}(p)$ and $q_x(\theta;p)$ are continuous in p, we have that for any fixed $s \in \mathbb{Z}^+$, the partial sum

$$\sum_{x\in\mathcal{X}\setminus\mathcal{X}_s^+} \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p) f_{x,\theta}(z_n^\star;p)$$

is continuous in p. On the other hand, we have $f_{x,\theta}(z^\star;p) \leqslant d_x^+ \mu_x q_x(\theta;p)$. Let C be a common upper bound for all $\bar{\ell}_{x,\theta}$. We thus have that

$$\sum_{x\in\mathcal{X}_s^+} \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta}(p) f_{x,\theta}(z_n^\star;p) \leqslant C \sum_{x\in\mathcal{X}_s^+} d_x^+ \mu_x,$$

which goes to zero as $s \to \infty$ by Assumption 3.4. Hence $f_\Gamma(z^\star;p)$ is continuous in p and combining with the continuity of the inverse demand function g, it follows that $\phi(p)$ is continuous in p. The same argument for ϕ_n lead to the continuity of ϕ_n on p, given the continuity of $z_n^\star(\cdot)$.

We next consider the case when $\bar{p} \in (p_{\min}, p_0)$ and there exists some small $\epsilon > 0$ such that $\phi(\bar{p} + \epsilon) > 0$ and $\phi(\bar{p} - \epsilon) < 0$. Notice that ϕ_n converges uniformly to ϕ since $f_\Gamma^{(n)}(z_n^\star(p); p)$ converges uniformly to $f_\Gamma(z^\star(p); p)$ on $[p_{\min}, p_0]$. So we have that for n large enough, $\phi_n(\bar{p} + \epsilon) > 0$ and $\phi_n(\bar{p} - \epsilon) < 0$. We can choose $\epsilon < \delta$ such that the local continuity of $z^\star(\cdot)$ and all $z_n^\star(\cdot)$ hold on $(\bar{p} - \epsilon, \bar{p} + \epsilon)$. By taking ϵ arbitrarily small, we conclude that $\bar{p}_n \to \bar{p}$ as $n \to \infty$. A similar argument gives the conclusion for $\bar{p} = p_{\min}$. This completes the proof of Lemma 3.25. ☐

3.6 Extension to Multiple Illiquid Assets

In this section, we expand our model to accommodate a financial network setup featuring multiple types of illiquid assets. We then state central limit theorems for a default cascade with fire sales within this setup.

Model. We consider K different illiquid assets $[K] := \{1, 2, \ldots, K\}$. Every institution holds a portfolio of illiquid assets $\boldsymbol{\gamma}_i = (\gamma_{i,1}, \ldots, \gamma_{i,K})^T$. We denote the average assets holdings by the vector $\bar{\boldsymbol{\gamma}} = (\bar{\gamma}_1, \ldots, \bar{\gamma}_K)^T$. For the initial price vector $\mathbf{p}_0 = (p_{0,1}, \ldots, p_{0,K})^T$ of the illiquid assets and given $\mathbf{p}_{\min} := (p_{\min,1}, \ldots, p_{\min,K})^T \leqslant \boldsymbol{p}_0$, we assume that there exists an exogenously given positive continuous inverse demand function for the multiple illiquid assets

$$\mathbf{g} := (g_1, \ldots, g_K)^T : \ [\mathbf{0}, \bar{\gamma}] \to [\mathbf{p}_{min}, \mathbf{p}_0],$$

with $g_k : [0, \bar{\gamma}_k] \to [p_{\min,k}, p_{0,k}]$, which satisfies Assumption 3.1, i.e.,

(i) $\mathbf{g}(\mathbf{0}) = \boldsymbol{p}_0$ (in absence of liquidations the price is given exogenously by $\boldsymbol{p}_0$).

(ii) For all $k \in [K], g_k(x) \in \mathcal{C}^1$ and it is a non-increasing function of $x \in [0, \bar{\gamma}_k]$ (the price is non-increasing with the average excess supply $\boldsymbol{x}$).

(iii) $\boldsymbol{g}(\bar{\gamma}) = \boldsymbol{p}_{\min} > 0$ (the price when the total illiquid asset holdings of the banks are sold is bounded from below by $\boldsymbol{p}_{\min} > 0$).

Similarly, for a given shock scenario $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n) \in [0, 1]^n$ and a given price $\boldsymbol{p} \in \mathbb{R}_+^K$ of the illiquid asset, we say that the bank i is ***p**-fundamentally insolvent* if its capital, after the shock and under price $\boldsymbol{p}$ of illiquid assets, is negative. We let the set of $\boldsymbol{p}$-fundamental defaults

$$\mathcal{D}_0(\boldsymbol{\epsilon}; \boldsymbol{p}) = \{i \in [n] : c_i(\epsilon_i; \boldsymbol{p}) < 0\}.$$

We next replace p by $\boldsymbol{p}$ for all definitions and assumptions of Section 4.2 and Section 3.3. In particular, for a given price $\boldsymbol{p}$, the default threshold distribution is now $q_x(\theta; \boldsymbol{p})$ for all $x \in \mathcal{X}$ and $\theta \in \{1, \ldots, d_x^+\} \cup \{\infty\}$.

For each asset $k \in [K]$, we let $\{L^{(i)}_{x,\theta,k}(\mathbf{p})\}_{i=1}^{\infty}$ be i.i.d. positive bounded random variables with common distribution $F_{x,\theta,k}(.;\mathbf{p})$, which has expectation $\bar{\ell}_{x,\theta,k}(\mathbf{p})$ and variance $\varsigma^2_{x,\theta,k}(\mathbf{p})$ under price $p \in [p_{\min}, p_0]$ for illiquid asset, for all $x \in \mathcal{X}$ and $\theta \in \{1, \ldots, d^+_x\} \cup \{\infty\}$.

Similarly to Assumption 3.4, we assume that the mean $\bar{\ell}_{x,\theta,k}(\mathbf{p})$ and variance $\varsigma^2_{x,\theta,k}(\mathbf{p})$ of sold shares for each liquidation are both continuous in $\mathbf{p}$ (on each p_k), for all $x \in \mathcal{X}$ and θ.

We recall that $L^{(i)}_{x,\theta,k}(\mathbf{p})$ denotes the units of k-th illiquid asset sold at i-th incoming default to institutions with type x and threshold θ. Further, $L^{(i)}_{x,\infty,k}(\mathbf{p})$ denotes the units of k-th illiquid asset sold at i-th incoming default to institutions with type x who never defaults.

For $k \in [K]$, the total sold shares of k-th illiquid asset at time t is given by (for price $\mathbf{p}$)

$$\Gamma^{(n)}_k(t;\mathbf{p}) := \sum_{x\in\mathcal{X}} \Big(\bar{\gamma}_{x,k} D^{(n)}_{x,0}(\mathbf{p}) + \sum_{\theta=1}^{d^+_x} Y^{(n)}_{x,\theta,k}(t;\mathbf{p}), + Y^{(n)}_{x,\infty,k}(t;\mathbf{p}) \Big), \tag{3.42}$$

where

$$Y^{(n)}_{x,\theta,k}(t;\mathbf{p}) := \sum_{i=1}^{I^{(n)}_{x,\theta}(t;p)} L^{(i)}_{x,\theta,k}(\mathbf{p}) \quad \text{and} \quad Y^{(n)}_{x,\infty,k}(t;\mathbf{p}) := \sum_{i=1}^{I^{(n)}_{x,\infty}(t;p)} L^{(i)}_{x,\infty,k}(\mathbf{p}). \tag{3.43}$$

The final shares of illiquid assets which have been sold under price $\mathbf{p}$ will be

$$\mathbf{\Gamma}^{(n)}(\tau^\star_n(\mathbf{p});\mathbf{p}) = \big(\Gamma^{(n)}_1(\tau^\star_n(\mathbf{p});\mathbf{p}), \ldots, \Gamma^{(n)}_K(\tau^\star_n(\mathbf{p});\mathbf{p}) \big)^T,$$

where $\Gamma^{(n)}_k(\tau^\star_n(\mathbf{p});\mathbf{p})$ denotes the final sold shares of illiquid asset k under price $\mathbf{p}$.

We next set the prices given by inverse demand function as

$$\boldsymbol{\kappa}_n(\mathbf{p}) := \boldsymbol{g}(\mathbf{\Gamma}^{(n)}(\tau^\star_n(\mathbf{p});\mathbf{p})/n).$$

Similarly, we define the *equilibrium prices* of the illiquid assets as

$$\mathbf{p}^\star_n = \sup\{\mathbf{p} \in [\mathbf{p}_{\min}; \mathbf{p}_0] : \mathbf{p} \leqslant \boldsymbol{\kappa}_n(\mathbf{p})\}, \tag{3.44}$$

where we take the supremum according to the K-dimensional Euclidean distance from $\mathbf{0}$.

Central Limit Theorems. We now discuss how the central limit theorem results from Section 3.3.2 can be extended to encompass multiple illiquid assets in the financial system. Following this, extending the other limit theorems (the law of large numbers) should be straightforward.

For each asset $k \in [K]$ and $z \in [0,1]$, we define

$$f_{\Gamma,k}(z;\mathbf{p}) := \sum_{x\in\mathcal{X}} \mu_x \Big(\bar{\gamma}_{x,k} q_x(0;\mathbf{p}) + \sum_{\theta=1}^{d^+_x} \bar{\ell}_{x,\theta,k}(\mathbf{p}) f_{x,\theta,k}(z;\mathbf{p}) + \bar{\ell}_{x,\infty,k}(\mathbf{p}) f_{x,\infty,k}(z;\mathbf{p}) \Big),$$

$$f^{(n)}_{\Gamma,k}(z;\mathbf{p}) := \sum_{x\in\mathcal{X}} \mu^{(n)}_x \Big(\bar{\gamma}_{x,k} q^{(n)}_x(0;\mathbf{p}) + \sum_{\theta=1}^{d^+_x} \bar{\ell}_{x,\theta,k}(\mathbf{p}) f^{(n)}_{x,\theta,k}(z;\mathbf{p}) + \bar{\ell}_{x,\infty,k}(\mathbf{p}) f^{(n)}_{x,\infty,k}(z;\mathbf{p}) \Big),$$

where

$$f_{x,\theta}(z;\mathbf{p}) := \mu_x q_x(\theta;\mathbf{p})\big(\theta - \sum_{\ell=d_x^+-\theta+1}^{d_x^+} \beta(d_x^+, z, \ell)\big), \quad f_{x,\infty}(z;\mathbf{p}) := (1-z)\mu_x q_x(\infty;\mathbf{p}) d_x^+,$$
$$f_{x,\theta}^{(n)}(z;\mathbf{p}) := \mu_x^{(n)} q_x^{(n)}(\theta;\mathbf{p})\big(\theta - \sum_{\ell=d_x^+-\theta+1}^{d_x^+} \beta(d_x^+, z, \ell)\big), \quad f_{x,\infty}^{(n)}(z;\mathbf{p}) := (1-z)\mu_x^{(n)} q_x^{(n)}(\infty;\mathbf{p}) d_x^+.$$

We also set the vectors

$$\boldsymbol{f}_\Gamma(z;\mathbf{p}) = \big(f_{\Gamma,1}(z;\mathbf{p}),\ldots,f_{\Gamma,K}(z;\mathbf{p})\big)^T \quad \text{and} \quad \boldsymbol{f}_\Gamma^{(n)}(z;\mathbf{p}) = \big(f_{\Gamma,1}^{(n)}(z;\mathbf{p}),\ldots,f_{\Gamma,K}^{(n)}(z;\mathbf{p})\big)^T.$$

Note that for any $k \in [K]$, the total sold shares for asset k, i.e., $\Gamma_k^{(n)}$, has the same shape as that of Γ_n in the uni-asset case. Hence, it is not hard to generalize our limit theorem on Γ_n in Section 3.3.2 to the multi-type case under the same assumptions. In particular, the following two theorems hold (under Assumption 3.5b for degree sequences), by systematically replacing p by $\mathbf{p}$ and considering each asset separately.

Theorem 3.26. *Let $\tau_n \leqslant \tau_n^\star(\mathbf{p})$ be a stopping time such that $\tau_n \xrightarrow{p} t_0$ for some $t_0 > 0$. Then for any fixed $k \in [K]; \mathbf{p} \in [\mathbf{p}_{\min}; \mathbf{p}_0]$ and $t > 0$, as $n \to \infty$,*

$$n^{-1/2}\big(\Gamma_k^{(n)}(t \wedge \tau_n;\mathbf{p}) - n\widehat{f}_{\Gamma,k}^{(n)}(t \wedge \tau_n;\mathbf{p})\big) \xrightarrow{d} \mathcal{Z}_{\Gamma,k}(t \wedge t_0;\mathbf{p}), \tag{3.45}$$

where $\mathcal{Z}_{\Gamma,k}(t;\mathbf{p})$ is a Gaussian random variable with mean 0 and variance

$$\Psi_k(t;\mathbf{p}) := \operatorname{Var}(\mathcal{Z}_{\Gamma,k}(t;\mathbf{p})),$$

where the form of $\Psi_k(t;\mathbf{p})$ is given by (3.48).

We also have the following theorem for the asymptotic normality of the final total sold shares.

Theorem 3.27. *Let $t^\star(\mathbf{p}) := -\ln z^\star(\mathbf{p})$. For any fixed $\mathbf{p} \in [\mathbf{p}_{\min}; \mathbf{p}_0]$, as $n \to \infty$, the final total sold shares for asset $k \in [K]$ satisfies:*

(i) If $z^\star(\mathbf{p}) = 0$ then asymptotically almost all institutions default after shock and (as $n \to \infty$)

$$\frac{\Gamma_k^{(n)}(\tau_n^\star; p)}{n} \xrightarrow{p} \sum_{x\in\mathcal{X}} \mu_x \Big(\bar{\gamma}_{x,k} q_x(0;\mathbf{p}) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta,k}(\mathbf{p})\theta q_x(\theta;\mathbf{p})\Big).$$

(ii) If $z^\star(\mathbf{p}) \in (0,1]$ and $z^\star(\mathbf{p})$ is a stable solution, i.e., $\alpha(\mathbf{p}) := f_W^1(z^\star(\mathbf{p});\mathbf{p}) > 0$, then

$$n^{-1/2}\big(\Gamma_k^{(n)}(\tau_n^\star;\mathbf{p}) - n f_{\Gamma,k}^{(n)}(z_n^\star(p);p)\big) \xrightarrow{d} \mathcal{Z}_{\Gamma,k}(t^\star(\mathbf{p});\mathbf{p}) - \alpha(\mathbf{p})^{-1} f_{\Gamma,k}^1(z^\star(\mathbf{p});\mathbf{p})\mathcal{Z}_W(t^\star(\mathbf{p})),$$

where $f_{\Gamma,k}^1$ denotes the partial derivative of $f_{\Gamma,k}$ with respect to the first variate z.

We next show a central limit theorem for the price $\boldsymbol{\kappa}_n(\mathbf{p}) := \boldsymbol{g}(\boldsymbol{\Gamma}^{(n)}(\tau_n^\star(\mathbf{p}); \mathbf{p})/n)$. We denote the vectors

$$\boldsymbol{\mathcal{Z}}_\Gamma(t; \mathbf{p}) := \big(\mathcal{Z}_{\Gamma,1}(t; \mathbf{p}), \ldots, \mathcal{Z}_{\Gamma,K}(t; \mathbf{p})\big)^T \quad \text{and} \quad \boldsymbol{f}_\Gamma^1(z; \mathbf{p}) = \big(f_{\Gamma,1}^1(z; \mathbf{p}), \ldots, f_{\Gamma,K}^1(z; \mathbf{p})\big)^T.$$

Theorem 3.28. *Let $t^\star(\mathbf{p}) := -\ln z^\star(\mathbf{p})$. For any $\mathbf{p} \in [\mathbf{p}_{\min}; \mathbf{p}_0]$ fixed and as $n \to \infty$, the price $\boldsymbol{\kappa}_n(p)$ given by the inverse demand function satisfies:*

(i) If $z^\star(\mathbf{p}) = 0$ then asymptotically almost all institutions default after shock and

$$\boldsymbol{\kappa}_n(p) \xrightarrow{p} \boldsymbol{g}\Big(\bar{\boldsymbol{\Gamma}}(\mathbf{p})\Big),$$

where $\bar{\boldsymbol{\Gamma}}(\mathbf{p}) := \big(\bar{\Gamma}_1(\mathbf{p}), \ldots, \bar{\Gamma}_K(\mathbf{p})\big)^T$ is given by setting, for all $k \in [K]$,

$$\bar{\Gamma}_k(\mathbf{p}) := \sum_{x \in \mathcal{X}} \mu_x\big(\bar{\gamma}_{x,k} q_x(0; \mathbf{p}) + \sum_{\theta=1}^{d_x^+} \bar{\ell}_{x,\theta,k}(\mathbf{p})\theta q_x(\theta; \mathbf{p})\big).$$

(ii) If $z^\star(\mathbf{p}) \in (0, 1]$ and $z^\star(\mathbf{p})$ is a stable solution, i.e., $\alpha(\mathbf{p}) := f_W^1(z^\star(\mathbf{p}); \mathbf{p}) > 0$, then

$$n^{-1/2}(\boldsymbol{\kappa}_n(\mathbf{p}) - \mathbf{g}(\mathbf{f}_\Gamma^{(n)}(z^\star(\mathbf{p}); \mathbf{p})) \xrightarrow{d} \mathbf{J}_\mathbf{g}\big(\mathbf{f}_\Gamma(z^\star(\mathbf{p}); \mathbf{p})\big)\Big[\boldsymbol{\mathcal{Z}}_\Gamma(t^\star(\mathbf{p}); \mathbf{p}) - \alpha(\mathbf{p})^{-1}\boldsymbol{f}_\Gamma^1(z^\star(\mathbf{p}); \mathbf{p})\mathcal{Z}_W(t^\star(\mathbf{p}); \mathbf{p})\Big],$$

where $\mathbf{J}_\mathbf{g}$ is the Jacobian matrix of $\mathbf{g}$.

Proof. The case $z^\star(\mathbf{p}) = 0$ is a direct consequent of point (i) of Theorem 3.27, since for all $k \in [K]$, g_k is continuous. Consider now the case when $z^\star(\mathbf{p}) \in (0, 1]$ and $z^\star(\mathbf{p})$ is a stable solution, i.e., $\alpha(\mathbf{p}) := f_W^1(z^\star(\mathbf{p}); \mathbf{p}) > 0$. Since the liquidations are independent for different types of assets, we have as a consequent of point (ii) of Theorem 3.16 that for all $k \in [K]$, $\Gamma_k^{(n)}(\tau_n^\star; \mathbf{p})$ is asymptotically normal and

$$n^{1/2}(\Gamma_k^{(n)}(\tau_n^\star; \mathbf{p})/n - f_{\Gamma,k}^{(n)}(z_n^\star; \mathbf{p})) \xrightarrow{d} \mathcal{Z}_k(t^\star; \mathbf{p}) - \alpha^{-1} f'_{\Gamma,k}(z^\star; \mathbf{p})\mathcal{Z}_W(t^\star; \mathbf{p}). \tag{3.46}$$

By a similar argument, we have that $z_n^\star(\mathbf{p}) \to z^\star(\mathbf{p})$ in probability and, for all $k \in [K]$,

$$f_{\Gamma,k}^{(n)}(z_n^\star; \mathbf{p}) \xrightarrow{p} f_{\Gamma,k}(z^\star; \mathbf{p}),$$

as $n \to \infty$, for any fixed $\mathbf{p} \in [\mathbf{p}_{min}, \mathbf{p}_0]$. Further, since the inverse demand function $\mathbf{g}$ is $\mathcal{C}^1$, we have for all $k \in [K]$, as $n \to \infty$,

$$g'_k \circ f_{\Gamma,k}^{(n)}(z_n^\star; \mathbf{p}) \xrightarrow{p} g'_k \circ f_{\Gamma,k}(z^\star; \mathbf{p}).$$

Hence, using the Mean-Value theorem, we have for all $k \in [K]$, there exists some $\boldsymbol{\Xi}_{k,n} := (\xi_{k,n}^{(1)}, \ldots, \xi_{k,n}^{(K)})$ converging to $\mathbf{f}_\Gamma(z^\star; \mathbf{p}) = (f_{\Gamma,1}(z^\star; \mathbf{p}), \ldots, f_{\Gamma,K}(z^\star; \mathbf{p}))$ such that

$$g_k(\boldsymbol{\Gamma}^{(n)}(\tau_n^\star; \mathbf{p})/n) - g_k(\boldsymbol{f}_\Gamma^{(n)}(z_n^\star; \mathbf{p})) = \nabla g_k(\boldsymbol{\Xi}_{k,n}) \cdot (\boldsymbol{\Gamma}^{(n)}(\tau_n^\star; \mathbf{p})/n - \mathbf{f}_\Gamma^{(n)}(z_n^\star; \mathbf{p})). \tag{3.47}$$

Multiply both side of (3.47) by $n^{1/2}$, we obtain

$$n^{1/2} g_k(\boldsymbol{\Gamma}^{(n)}(\tau_n^\star; \mathbf{p})/n) - g_k({}_\Gamma^{(n)}(z_n^\star; \mathbf{p})) = n^{1/2}\nabla g_k(\boldsymbol{\Xi}_{k,n}) \cdot (\boldsymbol{\Gamma}^{(n)}(\tau_n^\star; \mathbf{p})/n - \mathbf{f}_\Gamma^{(n)}(z_n^\star; \mathbf{p})).$$

Then by the asymptotic normality of point (ii) in theorem 3.27, we can generalize to our multidimensional case. The random vector $n^{1/2}(\mathbf{\Gamma}^{(n)}(\mathbf{p})/n - \mathbf{f}_\Gamma^{(n)}(z_n^\star, \mathbf{p}))$ converges in distribution to a centered Gaussian vector $\boldsymbol{\mathcal{Z}}_{end}(\mathbf{p}) := (\mathcal{Z}_{end}^{(1)}(\mathbf{p}), \ldots, \mathcal{Z}_{end}^{(K)}(\mathbf{p}))^T$, where

$$\mathcal{Z}_{end}^{(k)}(\mathbf{p}) := \mathcal{Z}_{\Gamma,k}(t^\star; \mathbf{p}) - \alpha^{-1} f_{\Gamma,k}^1(z^\star; \mathbf{p}) \mathcal{Z}_W(t^\star; \mathbf{p}).$$

Using Slutsky's theorem, we obtain for all $k \in [K]$, as $n \to \infty$,

$$n^{1/2}(\kappa_n^{(k)}(\mathbf{p}) - g_k(\boldsymbol{f}_\Gamma^{(n)}(z_n^\star, \mathbf{p}))) \xrightarrow{d} \nabla g_k(\mathbf{f}_\Gamma(z^\star; \mathbf{p})) \cdot \boldsymbol{\mathcal{Z}}_{end}(\mathbf{p}).$$

$\square$

We now state a central limit theorem for the equilibrium price after shock, defined by Equation (3.44). Define

$$\bar{\mathbf{p}} := \sup\{\mathbf{p} \in [\mathbf{p}_{min}, \mathbf{p}_0] : \mathbf{p} \leqslant \mathbf{g}(\mathbf{f}_\Gamma(z^\star(\mathbf{p}); \mathbf{p})\},$$

and correspondingly for the network of size n, we set

$$\bar{\mathbf{p}}_n := \sup\{\mathbf{p} \in [\mathbf{p}_{min}, \mathbf{p}_0] : \mathbf{p} \leqslant \mathbf{g}(\mathbf{f}_\Gamma^{(n)}(z^\star(\mathbf{p}); \mathbf{p})\},$$

We say that $\bar{\mathbf{p}}$ is a *stable* fixed point solution if either $\mathbf{p} = \mathbf{p}_{\min}$ or, $\mathbf{p} \in (\mathbf{p}_{\min}; \mathbf{p}_0]$ and there exists some $\boldsymbol{\epsilon} > 0$ such that $\mathbf{p} < \mathbf{g}(\mathbf{f}_\Gamma(z^\star(\mathbf{p}); \mathbf{p})$ for all $\mathbf{p} \in (\bar{\mathbf{p}} - \boldsymbol{\epsilon}; \bar{\mathbf{p}})$.

Let ∇f be the row vector of the gradient of f. For any function $f(z; \mathbf{p})$ and $k = 1, \ldots, K+1$, let $f^k(z; \mathbf{p})$ denote the partial derivative with respect to the k-th variate (z or p_{k-1}).

Theorem 3.29. *As $n \to \infty$, the equilibrium price satisfies:*

(i) If $z^\star(\bar{\mathbf{p}}) = 0$ and $\bar{\mathbf{p}}$ is a stable solution, then the equilibrium price converges to $\mathbf{p}_n^\star \xrightarrow{p} \bar{\mathbf{p}}$ and $\bar{\mathbf{p}}$ is the largest solution of the fixed point equation

$$\mathbf{p} = \mathbf{g}(\bar{\mathbf{\Gamma}}(\mathbf{p})),$$

where $\bar{\mathbf{\Gamma}}(\bar{p})$ is the same vector as defined in Theorem 3.28.

(ii) If $z^\star(\bar{\mathbf{p}}) \in (0, 1]$ is a stable solution of $f_W(z; \bar{\mathbf{p}}) = 0$, i.e., $\alpha(\bar{\mathbf{p}}) := f_W^1(z^\star(\bar{\mathbf{p}}); \bar{\mathbf{p}}) > 0$, and $\bar{\mathbf{p}}$ is a stable fixed point solution in $(\mathbf{p}_{\min}; \mathbf{p}_0)$, then

$$n^{-1/2}(\mathbf{p}_n^\star - \bar{\mathbf{p}}_n) \xrightarrow{d} -(\mathbf{I}_{K\times K} - \mathbf{A} \cdot \mathbf{B})^{-1} \boldsymbol{\mathcal{Z}}_V(\bar{\mathbf{p}}),$$

if the matrix $\mathbf{I}_{K\times K} - \mathbf{A} \cdot \mathbf{B}$ is non-singular, where $\mathbf{I}_{K\times K}$ is the $K \times K$ identity matrix, $\mathbf{A} = \mathbf{J}_{\mathbf{g}}\big(\mathbf{f}_\Gamma(z^\star(\mathbf{p}); \mathbf{p})\big)$ is the Jacobian matrix, $\mathbf{B}$ is also a $K \times K$ matrix with entry

$$\mathbf{B}_{ij} := -f_{\Gamma,i}^1(z^\star(\bar{\mathbf{p}}); \bar{\mathbf{p}}) \alpha(\bar{\mathbf{p}})^{-1} f_W^{j+1}(z^\star(\bar{\mathbf{p}}); \bar{\mathbf{p}}) + f_{\Gamma,i}^{j+1}(z^\star(\bar{\mathbf{p}}); \bar{\mathbf{p}})$$

for all $i, j \in [K]$, and $\boldsymbol{\mathcal{Z}}_V(\bar{\mathbf{p}}) := \big(\mathcal{Z}_{V,1}(\bar{\mathbf{p}}), \ldots, \mathcal{Z}_{V,K}(\bar{\mathbf{p}})\big)^T$ with (for $k \in [K]$)

$$\mathcal{Z}_{V,k}(\bar{\mathbf{p}}) := -\nabla g_k(\mathbf{f}_\Gamma(z^\star(\bar{\mathbf{p}}); \bar{\mathbf{p}}))\Big(\mathcal{Z}_{\Gamma,k}(t^\star(\bar{\mathbf{p}}); \bar{\mathbf{p}}) - \alpha(\bar{\mathbf{p}})^{-1} f_{\Gamma,k}^1(z^\star(\bar{\mathbf{p}}); \bar{\mathbf{p}}) \mathcal{Z}_W(t^\star(\bar{\mathbf{p}}); \bar{\mathbf{p}})\Big)$$

is a centered Gaussian random variable with mean 0.

Proof. We first provide the variance function of $\mathcal{Z}_{\Gamma,k}(t;\mathbf{p})$. Proceeding as in Section 3.5.6, one can show that for all $k \in [K]$, $\Psi_k(t;\mathbf{p})$ has the same structure as that in the uni-asset case. By replacing the corresponding mean $\bar{\ell}_{x,\theta,k}$, variance $\varsigma^2_{x,\theta,k}$ and $\bar{\gamma}_{x,k}$ for each type k, we get the variance function for $\mathcal{Z}_{\Gamma,k}(t;\mathbf{p})$, i.e.,

$$\begin{aligned}\Psi_k(t;\mathbf{p}) = \sum_{x\in\mathcal{X}}\big(\sum_{\theta=1}^{d_x^+}\Psi_{x,\theta,k}(t;\mathbf{p}) + \Psi_{x,\infty,k}(t;\mathbf{p}) + \bar{\gamma}^2_{x,k}\psi_{x,0,0}(\mathbf{p})\big) + \sum_{x\in\mathcal{X}}\sum_{\theta_1,\theta_2=1}^{d_x^+}\sigma^I_{x,\theta_1,\theta_2}(e^{-t};\mathbf{p})\\ + \sum_{x\in\mathcal{X}}\big(2\sum_{\theta=1}^{d_x^+}\sigma^I_{x,\theta,\infty}(e^{-t};\mathbf{p}) + 2\sum_{\theta=1}^{d_x^+}\psi_{x,\theta,0}(\mathbf{p})\sum_{s=d_x^+-\theta+1}^{d_x^+}\beta(d_x^+,e^{-t},s)\big)\\ + \sum_{x\in\mathcal{X}}\psi_{x,\infty,0}(\mathbf{p})\sum_{s=1}^{d_x^+}\beta(d_x^+,e^{-t},s),\end{aligned} \tag{3.48}$$

where for all $\theta \in \{1,\dots,d_x^+\}\cup\{\infty\}$

$$\Psi_{x,\theta,k}(t;\mathbf{p}) := \widehat{f}_{x,\theta}(t;\mathbf{p})\varsigma^2_{x,\theta,k}(\mathbf{p}) + \bar{\ell}^2_{x,\theta,k}(\mathbf{p})\sigma^I_{x,\theta,\theta}(e^{-t};\mathbf{p}).$$

We now proceed with the proof of Theorem 3.29. The case $z^\star(\mathbf{p}) = 0$ is a direct generalization of the corresponding situation in Theorem 3.19 and can be proved by a similar argument, using Theorem 3.28. Consider now the case when $z^\star(\mathbf{p}) \in (0,1]$ and $z^\star(\mathbf{p})$ is a stable solution, i.e., $\alpha(\mathbf{p}) := f^1_W(z^\star(\mathbf{p});\mathbf{p}) > 0$. First of all, Lemma 3.25 can be generalized to the multi-dimensional case and shows that $\bar{\mathbf{p}}_n \to \bar{\mathbf{p}}$. Further, we also have that

$$\boldsymbol{\mathcal{Z}}_V(\mathbf{p}_n) \xrightarrow{d} \boldsymbol{\mathcal{Z}}_V(\mathbf{p}), \tag{3.49}$$

for any sequence $\{\mathbf{p}_n\}_n$ converging to $\mathbf{p}$ as $n\to\infty$. Let us denote

$$\Phi_n^{(k)}(\mathbf{p}) := p_k - g_k(\mathbf{\Gamma}^{(n)}(\tau_n^\star;\mathbf{p})/n).$$

We use the Skorokhod representation theorem. All the convergence results of Theorem 3.28, $\mathbf{p}_n^\star \to \bar{\mathbf{p}}$ and (3.49) hold a.s., by changing the probability space. Then we can write

$$\begin{aligned}\Phi_n^{(k)}(\mathbf{p}_n^\star) &= \phi_n^{(k)}(\mathbf{p}_n^\star) + n^{-1/2}\mathcal{Z}_{V,k}(\mathbf{p}_n^\star) + o(n^{-1/2})\\ &= \phi_n^{(k)}(\mathbf{p}_n^\star) + n^{-1/2}\mathcal{Z}_{V,k}(\bar{\mathbf{p}}) + o(n^{-1/2}),\end{aligned} \tag{3.50}$$

where the second equality holds because we have a.s. $\mathcal{Z}_{V,k}(\mathbf{p}_n^\star) \to \mathcal{Z}_{V,k}(\bar{\mathbf{p}})$. Notice also that for all $k\in[K]$, $\Phi_n^{(k)}(\mathbf{p}_n^\star) = 0$, and we have

$$\phi_n^{(k)}(\mathbf{p}_n^\star) = -n^{-1/2}\mathcal{Z}_{V,k}(\bar{\mathbf{p}}) + o(n^{-1/2}). \tag{3.51}$$

We proceed to approximate the difference between $\mathbf{p}_n^\star$ and $\bar{\mathbf{p}}_n$ by the Mean-Value theorem. The arguments are similar to the uni-asset case; we thus just highlight the difference from the one asset situation. We denote by $o(\mathbf{1})$ the K-column vector of $o(1)$. For all $k\in[K]$, we have

$$\begin{aligned}&g_k(\mathbf{f}_\Gamma^{(n)}(z_n^\star(\mathbf{p}_n^\star);\mathbf{p}_n^\star)) - g_k(\mathbf{f}_\Gamma^{(n)}(z_n^\star(\bar{\mathbf{p}}_n);\bar{\mathbf{p}}_n))\\ &\quad= (\nabla g_k(\mathbf{f}_\Gamma(z^\star(\bar{\mathbf{p}});\bar{\mathbf{p}})) + o(\mathbf{1}))(\mathbf{f}_\Gamma^{(n)}(z_n^\star(\mathbf{p}_n^\star);\mathbf{p}_n^\star) - \mathbf{f}_\Gamma^{(n)}(z_n^\star(\bar{\mathbf{p}}_n);\bar{\mathbf{p}}_n)).\end{aligned} \tag{3.52}$$

Next we analyze $\mathbf{f}_\Gamma^{(n)}(z_n^\star(\mathbf{p}_n^\star);\mathbf{p}_n^\star) - \mathbf{f}_\Gamma^{(n)}(z_n^\star(\bar{\mathbf{p}}_n);\bar{\mathbf{p}}_n)$. By the Mean-Value theorem, we have for all $k \in [K]$,

$$
\begin{aligned}
f_k^{(n)}(z_n^\star(\mathbf{p}_n^\star);\mathbf{p}_n^\star) - f_k^{(n)}(z_n^\star(\bar{\mathbf{p}}_n);\bar{\mathbf{p}}_n) =& (f_k^1(z^\star(\bar{\mathbf{p}});\bar{\mathbf{p}}) + o(1))(z_n^\star(\mathbf{p}_n^\star) - z_n^\star(\bar{\mathbf{p}}_n)) \\
&+ (\nabla^{(2)} f_k(z^\star(\bar{\mathbf{p}});\bar{\mathbf{p}}) + o(\mathbf{1})) \cdot (\mathbf{p}_n^\star - \bar{\mathbf{p}}_n).
\end{aligned} \tag{3.53}
$$

We next approximate $z_n^\star(\mathbf{p}_n^\star) - z_n^\star(\bar{\mathbf{p}}_n)$. Notice that $f_W^{(n)}(z_n^\star(\mathbf{p});\mathbf{p}) = 0$ for any $\mathbf{p} \in [\mathbf{p}_{min}, \mathbf{p}_0]$. Using again the Mean-Value theorem, we have the following two equations

$$
f_W^{(n)}(z_n^\star(\mathbf{p}_n^\star);\mathbf{p}_n^\star) - f_W^{(n)}(z_n^\star(\bar{\mathbf{p}}_n);\mathbf{p}_n^\star) = f_W^{1,(n)}(\xi_n^z;\mathbf{p}_n^\star)(z_n^\star(\mathbf{p}_n^\star) - z_n^\star(\bar{\mathbf{p}}_n)),
$$

and,

$$
f_W^{(n)}(z_n^\star(\bar{\mathbf{p}}_n);\mathbf{p}_n^\star) - f_W^{(n)}(z_n^\star(\bar{\mathbf{p}}_n);\bar{\mathbf{p}}_n) = \nabla^{(2)} f_W^{(n)}(z_n^\star(\bar{\mathbf{p}}_n),\boldsymbol{\zeta}_n) \cdot (\mathbf{p}_n^\star - \bar{\mathbf{p}}_n),
$$

where $\xi_n^z \to z^\star(\bar{\mathbf{p}})$ a.s., $\boldsymbol{\zeta}_n \to \bar{\mathbf{p}}$ a.s. and the notation $\nabla^{(2)}$ is defined by setting

$$
\nabla^{(2)} F(z;\mathbf{p}) = (F^2(z;\mathbf{p}), \dots, F^{K+1}(z;\mathbf{p})),
$$

Then by the above two equations we have

$$
z_n^\star(\mathbf{p}_n^\star) - z_n^\star(\bar{\mathbf{p}}_n) = -((f_W^1(z^\star(\bar{\mathbf{p}});\bar{\mathbf{p}}))^{-1} + o(\mathbf{1}))(\nabla^{(2)} f_W(z^\star(\bar{\mathbf{p}});\bar{\mathbf{p}}) + o(\mathbf{1})) \cdot (\mathbf{p}_n^\star - \bar{\mathbf{p}}_n). \tag{3.54}
$$

Now by (3.53) and (3.54), we have that for all $k \in [K]$,

$$
\begin{aligned}
& f_{\Gamma,k}^{(n)}(z_n^\star(\mathbf{p}_n^\star);\mathbf{p}_n^\star) - f_{\Gamma,k}^{(n)}(z_n^\star(\bar{\mathbf{p}}_n);\bar{\mathbf{p}}_n) \\
&\quad = -f_k^1(z^\star(\bar{\mathbf{p}});\bar{\mathbf{p}})(f_W^1(z^\star(\bar{\mathbf{p}});\bar{\mathbf{p}}))^{-1}\nabla^{(2)} f_W(z^\star(\bar{\mathbf{p}});\bar{\mathbf{p}}) \cdot (\mathbf{p}_n^\star - \bar{\mathbf{p}}_n) \\
&\quad\quad + \nabla^{(2)} f_k(z^\star(\bar{\mathbf{p}});\bar{\mathbf{p}}) \cdot (\mathbf{p}_n^\star - \bar{\mathbf{p}}_n) + o(\mathbf{1}) \cdot (\mathbf{p}_n^\star - \bar{\mathbf{p}}_n) \\
&\quad = (\mathbf{B}_k + o(\mathbf{1})^T)(\mathbf{p}_n^\star - \bar{\mathbf{p}}_n),
\end{aligned} \tag{3.55}
$$

where $\mathbf{B}_k$ is the k-th row vector of $\mathbf{B}$. Hence by (3.52) and (3.55), for all $k \in [K]$, we can conclude

$$
\begin{aligned}
\phi_n^{(k)}(\mathbf{p}_n^\star) &= \phi_n^{(k)}(\mathbf{p}_n^\star) - \phi_n^{(k)}(\bar{\mathbf{p}}_n) \\
&= p_n^{\star(k)} - \bar{p}_n^{(k)} - (g_k(\mathbf{f}_\Gamma^{(n)}(z_n^\star(\mathbf{p}_n^\star);\mathbf{p}_n^\star)) - g_k(\mathbf{f}_\Gamma^{(n)}(z_n^\star(\bar{\mathbf{p}}_n);\bar{\mathbf{p}}_n))) \\
&= p_n^{\star(k)} - \bar{p}_n^{(k)} - (\mathbf{A}_k \cdot \mathbf{B} + o(\mathbf{1})^T)(\mathbf{p}_n^\star - \bar{\mathbf{p}}_n),
\end{aligned}
$$

where $\boldsymbol{A}_k$ is the k-th row vector of $\boldsymbol{A}$. Combining with Equation (3.51), it follows that

$$
\boldsymbol{\phi}_n(\mathbf{p}_n^\star) = (\mathbf{I}_{K\times K} - \mathbf{A}\cdot\mathbf{B} - o(\mathbf{1}_{K\times K}))(\mathbf{p}_n^\star - \bar{\mathbf{p}}_n) = -n^{-1/2}(\boldsymbol{\mathcal{Z}}_V(\bar{\mathbf{p}}) + o(\mathbf{1})).
$$

We therefore obtain that

$$
n^{1/2}(\mathbf{p}_n^\star - \bar{\mathbf{p}}_n) = -(\mathbf{I}_{K\times K} - \mathbf{A}\cdot\mathbf{B})^{-1}\boldsymbol{\mathcal{Z}}_V(\bar{\mathbf{p}}) + o(\mathbf{1}),
$$

provided that the matrix $\mathbf{I}_{K\times K} - \mathbf{A}\cdot\mathbf{B}$ is non-singular. □

3.7 Concluding Remarks

We have proposed a stochastic framework for quantifying the impact of a macroeconomic shock on the resilience of a banking network to fire sales and insolvency cascades. Our limit theorems provide quantitative evidence for the importance of monitoring fire sales and controlling indirect contagion in financial systems. We have quantified how price mediated contagion across institutions with common asset holding can worsen cascades of insolvencies in a heterogeneous financial network, during a financial crisis. Under suitable assumptions on the degree and threshold distributions, we have shown that the default cascade model can be transferred to a death process problem. This allows us to provide the limit theorems for a dynamic default cascade process with fire sales. We have stated various limit theorems regarding the total sold shares and the equilibrium price of illiquid assets in a stylized fire sales model. In particular, the equilibrium prices of illiquid assets have asymptotically Gaussian fluctuations. Additionally, we have established a link between the variance of these fluctuations and the characteristics of the financial network, as well as the inverse demand function.

In our numerical experiments, we investigated the effect of heterogeneity in network structure and price impact function on the final size of default cascade and fire sales loss. For a regular financial network, we found that for a small shock, the high connectivity network is more resilient. However, once the shock is large enough, the default propagates to a larger fraction of institutions due to its higher connectivity. On the other hand, the fire sales loss in the two financial networks with high and low connectivity are very close to each other. We also observed that financial networks with higher heterogeneity may have a smaller critical value for the shock beyond which a large fraction of institutions default, both with and without fire sales. However, for smaller shocks, the most heterogeneous network could be the least resilient. Additionally, we demonstrated the practical applicability of our central limit theorems by utilizing them to construct confidence intervals for the final fraction of defaults and fire sales loss in finite networks. These confidence intervals provide valuable insights into the uncertainty associated with the outcomes in such networks, allowing for a more robust assessment of systemic risk.

Our theoretical analysis provides valuable insights into financial stability and systemic risk. It highlights the importance of ensuring that a financial network can withstand large cascades under stress scenarios that put pressure on specific characteristics such as capital or liquidity reserves. To mitigate the risk of phase transitions and systemic instability, regulators may consider imposing higher capital requirements on financial institutions based on their classes. This would help prevent the occurrence of large cascades and enhance the resilience of the overall financial system. Additionally, the limit theorems we derived for heterogeneous financial networks offer interpretable and computationally efficient tools for regulators to assess systemic risk. By collecting data on network characteristics and utilizing these limit theorems, regulators can gain deeper insights into the potential risks and vulnerabilities within the financial system. This can inform their decision-making process and enable them to take appropriate measures to safeguard financial stability.

Several directions emerge from the current study.

- An important challenge for future research is to investigate the limitations and boundaries of central limit theorems in capturing extreme events and tail risk in financial networks, considering the potential deviations from Gaussian fluctuations in such scenarios.

- Another valuable area for future research lies in extending the central limit theorems to dynamic networks, where the network structure evolves over time, allowing for a more realistic representation of contagion processes in evolving financial systems.
- To provide a more comprehensive analysis of the interplay between market dynamics, financial network structure, and the overall stability of the system, a challenging direction for future research would be to endogenize the inverse demand function within the model.
- Another interesting avenue for future work is to endogenize the financial network payments. Currently, we have assumed exogenous interbank liabilities based on a specified distribution. However, allowing for endogenous determination of interbank payments could provide a more realistic representation of the complex dynamics within a financial network. This could involve incorporating feedback mechanisms between the financial health of institutions, their interbank exposures, and the resulting payment flows.
- Furthermore, there are other related issues that warrant further investigation. For instance, exploring the impact of regulatory policies and interventions on the resilience of financial networks and the occurrence of default cascades could provide valuable insights for policymakers. Additionally, studying the implications of network formation and evolution over time, as well as incorporating behavioral aspects and investor heterogeneity, can contribute to a more comprehensive understanding of systemic risk.

We leave these and related issues for future work.

Chapter 4

Ruin probabilities for Risk Processes in Stochastic Networks

This chapter is based on paper [5] in the publication list of Section 1.5.

Abstract. We study multidimensional Cramér-Lundberg risk processes where agents, located on a large sparse network, receive losses from their neighbors. To reduce the dimensionality of the problem, we introduce classification of agents according to an arbitrary countable set of types. The ruin of any agent triggers losses for all of its neighbors. We consider the case where the loss arrival process induced by the ensemble of ruined agents follows a Poisson process with general intensity function that scales with the network size. When the size of the network goes to infinity, we provide explicit ruin probabilities at the end of the loss propagation process for agents of any type. These limiting probabilities depend, in addition to the agents' types and the network structure, on the loss distribution and the loss arrival process. For a more complex risk processes on open networks, when in addition to the internal networked risk processes the agents receive losses from external users, we provide bounds on ruin probabilities.

Keywords: Risk processes, Cramér-Lundberg model, Ruin probabilities, Stochastic networks.

Contents

3.1 Introduction **91**
3.2 The Model **94**
3.2.1 Financial network 94
3.2.2 Default cascade 96
3.2.3 Node classification 97
3.2.4 A simple model of fire sales 99
3.3 Limit Theorems **101**
3.3.1 Asymptotic magnitude of default cascade with fire sales 102

3.3.2 Asymptotic normality of default cascade with fire sales 105
3.4 Numerical Experiments 109
3.4.1 Numerical set-up 110
3.4.2 Regular financial networks 111
3.4.3 Core-Periphery financial networks 113
3.4.4 Scale-Free financial networks 115
3.4.5 Asymptotic normality and confidence intervals 117
3.5 Proofs 121
3.5.1 Auxiliary lemmas 121
3.5.2 Proof of Theorem 3.7 124
3.5.3 Proof of Theorem 3.9 127
3.5.4 Proof of Theorem 3.11 127
3.5.5 Discussion on Lemma 3.6 and Lemma 3.13 128
3.5.6 Proof of Theorem 3.14 130
3.5.7 Proof of Theorem 3.16 136
3.5.8 Proof of Theorem 3.18 138
3.5.9 Proof of Theorem 3.19 139
3.6 Extension to Multiple Illiquid Assets 142
3.7 Concluding Remarks 149

4.1 Introduction

The classical compound risk process with Poisson claim arrivals, or the Cramér-Lundberg model ([100, 171]) has been extensively used in quantitative risk management, see e.g., [103, 173]. In this model, the aggregate capital of an insurer who starts with initial capital γ, premium rate α and (loss) claim sizes L_k is given by the following spectrally negative compound Poisson process

$$C(t) = \gamma + \alpha t - \sum_{k=1}^{\mathcal{N}(t)} L_k, \tag{4.1}$$

where $L_k, k \in \mathbb{N}$, are i.i.d. non-negative random variable following a distribution F with mean $\bar{L}$ and $\mathcal{N}(t)$ is a Poisson process with intensity $\beta > 0$ independent of L_k. The ruin time for the insurer with initial capital γ is defined by

$$\tau(\gamma) := \inf\{t \mid C(t) \leqslant 0\},$$

(with the convention that $\inf \varnothing = \infty$) and the central question is to find the ruin probability

$$\psi(\gamma) := \mathbb{P}(\tau(\gamma) < \infty).$$

It is known (see e.g. [35, 115]) that whenever $\beta\bar{L} > \alpha$, we have $\psi(\gamma) = 1$ for all $\gamma \in \mathbb{R}$ and whenever $\beta\bar{L} < \alpha$, the ruin probability can be computed using the famous Pollaczek–Khinchine formula as

$$\psi(\gamma) = \left(1 - \frac{\beta\bar{L}}{\alpha}\right) \sum_{k=0}^{\infty} \left(\frac{\beta\bar{L}}{\alpha}\right)^k \left(1 - \widehat{F}^{*k}(\gamma)\right), \tag{4.2}$$

where

$$\widehat{F}(\gamma) = \frac{1}{\bar{L}} \int_0^\gamma \big(1 - F(u)\big) du,$$

and the operator $(\cdot)^{*k}$ denotes the k-fold convolution.

We consider a (stochastic) large sparse network setting that replaces the standalone jump process in the classical model. In our model, losses of any firm do not occur via an exogenous Poisson process, but due to the ruin of its neighboring nodes (agents).

While the classical Cramér-Lundberg model and its generalizations such as Sparre Andersen model ([31]) have been a pillar for collective risk theory for many decades, ruin problems beyond one dimension remain challenging. Even numerical approximations are only available for some distributions, in low dimensions (see e.g., [39, 41]) and for some particular hierarchical structures. For example, [38] consider a central branch with one subsidiary which do not admit exact solution due to their complex dependent Sparre Andersen structure. The authors propose approximation techniques based on replacing the underlying structure with spectrally negative Markov additive processes. In [56, 158] the authors consider loss propagation in bipartite graphs. For exponential claims, the authors provide Pollaczek–Khinchine formulas for the summative ruin probability of a group of agents. One should note that this is not a model of agent interrelation, but a model in which insurers connect to 'objects' (external risks). As such, notions of losses that propagate from one ruined agent to another agent do not gave a correspondence in that model.

Network based models have been used to advance risk assessment in financial systems. An initial body of literature is concerned with economic questions such as the effects of network structure on financial contagion [1, 24, 93, 124, 143] or with considering and integrating variations of the distress propagation mechanism, see e.g., [58, 114, 132]. The question of control of financial contagion is posed in [26, 27], where authors introduce a link revealing filtration and adaptive bailouts mitigate the extent of contagion. Many of the initial models are static, in the sense that there is one snapshot of the network and an initial wave of defaults leads to a second wave of defaults and so on. The state of the network does not change over time. Rather, it is reassessed in rounds, in order to find a final set of defaults. Dynamic contagion is considered for example in [81, 119], where nodes are endowed with stochastic processes, usually jump-diffusions.

This chapter originates with the asymptotic analysis in [20]. Their results are purely static and their focus is to provide a condition of resilience under which the contagion does not spread to a strictly positive fraction of the agents. In Chapter 2, we provide steps for the asymptotic fraction of the ruined agents when nodes may have special ruin dynamics without growth; Unlike [20], the analysis is not based on the differential equation method. Instead, we generalize the law of large numbers for a model of default cascades in the configuration model. This allows us to go beyond the static model, and the approach could be used to provide central limit theorems. Our dynamic leads to a special case of loss intensity. All outgoing half-edges are assigned an exponential clock with parameter one, which determines when losses are revealed. The loss reveal intensity function is then equal to the number of remaining outgoing half-edges at time t. This model can be seen a variant of the notion of Parisian ruin: indeed, an agent could have become ruined if the incoming loss from a ruined neighbor was observed instantly. Instead, it is observed after an exponential time, during which the agents' capital increases. Therefore, it could well be that it withstands the loss at the time when it is revealed.

The closest literature to this chapter is [29]. They allow to linear growth at node level (proportional to the number of the node's links), while the loss reveal intensity is assumed to be constant equal to the size of network. The key feature that allowed the asymptotic analysis was the fact that each link carried a constant loss. This in turn, made their analysis simpler. Because losses can be arbitrary, the risk process does not evolve according to a given grid as in the case of constant losses. When a firm suffers a loss due to a neighbor failure it will moving to a lower level of value. Here, in presence of general losses, all possible value levels are coupled.

The remaining open question, posed in [29], is to allow for losses that come from a general distribution, as opposed to constant losses. In this chapter, we solve this open question, and in doing so we bridge the risk literature on multi-dimensional risk processes with the financial network literature, and use the asymptotic results of the latter to provide the approximations of the ruin probabilities.

Outline. In Section 4.2 we introduce the model of interconnected risk processes. The model is driven by a classification of nodes according to types, whereas the interconnections occur according to the configuration model. In Section 4.3 we state our main results, concerning the asymptotic fraction of ruined nodes. Section 4.4 provides the proofs. In Section 4.5 we outline a complex risk process driven by both exogenous individual external loss processes and an internal risk processes where losses propagate in the network. Section 4.6 concludes and proposes several open questions.

Notations. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If $c \in \mathbb{R}$ is a constant, we write $X_n \xrightarrow{p} c$ to denote that X_n converges in probability to c. Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers that tends to infinity as $n \to \infty$. We write $X_n = o_p(a_n)$, if $|X_n|/a_n \xrightarrow{p} 0$. We say that an event holds w.h.p. (with high probability), if it holds with probability tending to 1 as $n \to \infty$. For any function f defined on $\mathbb{R}^+ := [0, \infty)$, $||f||_{L^1}$ denotes the L^1 norm of f on $\mathbb{R}^+$. For any event or set A, we denote by A^c the complement of A. $\mathsf{Bin}(k, p)$ denotes a binomial distribution corresponding to the number of successes of a sequence of k independent Bernoulli trials each having probability of success p. We denote by $\mathbb{1}\{\mathcal{E}\}$ the indicator of an event $\mathcal{E}$, his is 1 if $\mathcal{E}$ holds and 0 otherwise. Let $\mathbb{R}_0^+$ denote the half line $[0, \infty)$. For $a, b \in \mathbb{R}$, we denote by $a \wedge b := \min\{a, b\}$.

4.2 The Model

4.2.1 Networked risk processes

Consider a set of n agents (e.g., firms, insurers, re-insurers, business lines, ...) denoted by $[n] := \{1, 2, \ldots, n\}$. Agents hold contractual obligations with each other. In our networked risk processes model, the interaction of the agents' capital processes occurs through the network of obligations. Upon the ruin of an agent, we consider that it will fail on their obligations and the neighboring agents will suffer losses due to this non fulfillment. Let $\mathcal{G}_n$ be a directed graph on $[n]$. For two agents $i, j \in [n]$, we write $i \to j$ when there is a directed edge (link) from i to j in $\mathcal{G}_n$, modeling the fact that i has a contractual obligation to j. Similar to the classical model (4.1), each agent i is endowed with an initial capital $\gamma_i > 0$, while $\alpha_i(t)$ is a continuous non-decreasing function describing the premium accumulation for agent i.

Here, we assume that the capital is affected by an initial exogenous proportional shock $\epsilon_i \in [0, 1]$ and δ_i represents the total value of claims held by end-users on agent i (deposits). It is then possible for an agent to fail after the exogenous shock if its initial capital is lower than the end-user claims, i.e., $C_i(0) := \gamma_i(1 - \epsilon_i) + \alpha_i(0) - \delta_i \leqslant 0$. The set of fundamentally ruined agents is thus

$$\mathcal{D}(0) := \{i \in [n] : C_i(0) \leqslant 0\}.$$

Ruined agents affect their neighbors through the network of obligations. The ruin time for agent $i \in [n]$ is $\tau_i := \inf\{t \mid C_i(t) \leqslant 0\}$ where we consider the following risk process for the capital of agent i with network interactions $\mathcal{G}_n$:

$$C_i(t) := \gamma_i(1 - \epsilon_i) + \alpha_i(t) - \delta_i - \sum_{j \in [n]: j \to i} L_{ji} \mathbb{1}\{\tau_j + T_{ji} \leqslant t\}. \tag{4.3}$$

Here we denote by T_{ji} the delay between firm j's ruin and the time when a neighbor i processes its losses from the unfulfilled obligations of j to i. At time $\tau_j + T_{ji}$, node j processes a loss L_{ji}. Similar to the classical model, we assume that the incoming losses of each agent $i \in [n]$ are i.i.d. random variables following some positive distribution F_i (potentially depending on the agent i characteristics). In order to compare with the classical model, we will assume that T_{ji} are i.i.d. exponentially distributed with some parameter $\beta > 0$, i.e., $T_{ji} \sim \mathrm{Exp}(\beta)$ for all $i, j \in [n]$.

Consequently, the set of ruined agents at time $t \geqslant 0$ will be given by

$$\mathcal{D}(t) := \{i \in [n] : C_i(t) \leqslant 0\} = \{i \in [n] : \tau_i \leqslant t\}.$$

We assume that agent i becomes inactive upon ruin, so that $C_i(t) = C_i(t \wedge \tau_i)$.

4.2.2 Node classification

We consider a classification of agents into a countable (finite or infinite) set of types $\mathcal{X}$. We denote by $x_i \in \mathcal{X}$ the type of agent $i \in [n]$.

Let us denote by $\mu_x^{(n)}$ the fraction of agents in class $x \in \mathcal{X}$ in the network $\mathcal{G}_n$. In order to study the asymptotics, it is natural to assume there is a limiting distribution of types.

Assumption 4.1. *We assume that for some probability distribution μ over $\mathcal{X}$ and independent of n, we have that $\mu_x^{(n)} \to \mu_x$, as $n \to \infty$, for all $x \in \mathcal{X}$.*

Since the type space is countable, we can assume without loss of generality that all agents of same type x (for all $x \in \mathcal{X}$) have the same number of outgoing links, denoted by d_x^-, and the same number of incoming links, denoted by d_x^+.

To reduce the dimensionality of the networked risk processes problem, we further assume that all parameters are type dependent. Namely, we assume that $\gamma_i = \gamma_x, \alpha_i(.) = \alpha_x(.)$ and $\delta_i = \delta_x$ for all agents $i \in [n]$ with $x_i = x$. Note that the type space can be made sufficiently large (but countable) to incorporate a wide variety of levels for these parameters.

In particular, shocks are assumed to be independent random variables with distribution function (cdf) $F_x^{(\epsilon)}$ and density function $f_x^{(\epsilon)}$ depending on the type of each agent. We then set

$$q_{x,0} := 1 - F_x^{(\epsilon)}\big(\frac{\gamma_x + \alpha_x(0) - \delta_x}{\gamma_x}\big), \tag{4.4}$$

which represents the (expected) fraction of initially ruined agents of type $x \in \mathcal{X}$.

The distribution of incoming losses for each agent are also assumed to be type dependent random variables. For all agents of type $x \in \mathcal{X}$, the loss distribution function (cdf) is denoted by F_x and the probability density function (pdf) is f_x. Thus we have $F_i = F_{x_i}$ for any agent $i \in [n]$.

Remark 4.1 (From loss distribution to threshold distribution)**.** *The model, as introduced, can be equivalent to a model of dynamic failure thresholds inferred from the loss distribution. These random thresholds measure how many ruined neighbors can an agent withstand before being ruined due to the incurred losses. For $x \in \mathcal{X}$, let ϵ_x be a random variable with distribution $F_x^{(\epsilon)}$ and $\{L_x^{(k)}\}_{k=1}^{\infty}$ be a set of i.i.d. positive continuous random variables with common cumulative distribution function (cdf) F_x. The threshold distribution function at time t is defined as $q_{x,0}(t) = q_{x,0}$,*

$$q_{x,1}(t) := \mathbb{P}\big(0 < \gamma_x(1-\epsilon_x) + \alpha_x(t) - \delta_x \leqslant L_x^{(1)}\big),$$

and for all $\theta \geqslant 2$,

$$q_{x,\theta}(t) := \mathbb{P}\big(L_x^{(1)} + \cdots + L_x^{(\theta-1)} < \gamma_x(1-\epsilon_x) + \alpha_x(t) - \delta_x \leqslant L_x^{(1)} + \cdots + L_x^{(\theta)}\big)$$

represents the probability at time t that an agent of type x is ruined after θ neighboring ruins. Since $\alpha_x(t)$ is a non-decreasing function of t, this threshold function will be (stochastically) non-decreasing. Note that when $\alpha_x(t) = 0$ over each class $x \in \mathcal{X}$, the results in Chapter 2 could be applied by using the above threshold distributions. Note that in these works, the threshold distributions are static, which makes the model simpler to study. Here the distributions change over time because there is time-dependent growth. The closest model is [29] which consider fixed losses and exponential inter-arrival times with fixed parameter. In this chapter, in order to study the general setup, we do not use these threshold distributions - which are given here only for comparison. Instead, our analysis relies on a sequence of random threshold times representing the times where thresholds hit subsequent levels.

We consider the risk processes in random network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ (the configuration model defined in Chapter 2), satisfying the following regularity condition on the average degrees.

Assumption 4.2. *We assume that, as $n \to \infty$, the first moment of degrees converges and is finite:*

$$\lambda^{(n)} := \sum_{x \in \mathcal{X}} d_x^+ \mu_x^{(n)} = \sum_{x \in \mathcal{X}} d_x^- \mu_x^{(n)} \overset{(as\ n\to\infty)}{\longrightarrow} \lambda := \sum_{x \in \mathcal{X}} d_x^+ \mu_x \in (0, \infty).$$

4.2.3 The loss reveal process

In order to study the risk processes in random network $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$, we construct the configuration model simultaneously as we run the ruin propagation model. Starting from the set of initially ruined agents, at each step, we only look at one interaction between a ruined agent and its counterparty in the configuration model. Note that the set of ruined agents either stays the same or augments with each such interaction. We first introduce some notations.

We denote by $\mathbf{D}^{(n)}(t)$ the set of ruined agents at time t and set $D^{(n)}(t) := |\mathbf{D}^{(n)}(t)|$. Similarly, $\mathbf{S}^{(n)}(t) = [n]/\mathbf{D}^{(n)}(t)$ denotes the set of solvent agents at time t and we set $S^{(n)}(t) = |\mathbf{S}^{(n)}(t)|$. For $x \in \mathcal{X}$, we denote by $\mathbf{S}_x^{(n)}(t)$ the set of all solvent agents in class x at time t and set $S_x^{(n)}(t) = |\mathbf{S}_x^{(n)}(t)|$. Moreover, for $x \in \mathcal{X}$ and $\theta = 0, \dots, d_x^+$, we denote by $\mathbf{S}_{x,\theta}^{(n)}(t)$ the set of all solvent agents in $\mathbf{S}_x^{(n)}(t)$ with exactly θ ruined incoming neighbors at time t. Set $S_{x,\theta}^{(n)}(t) = |\mathbf{S}_{x,\theta}^{(n)}(t)|$. We define similarly the sets $\mathbf{D}_x^{(n)}(t)$ and $\mathbf{D}_{x,\theta}^{(n)}(t)$ with corresponding sizes $D_x^{(n)}(t)$ and $D_{x,\theta}^{(n)}(t)$. We call all outgoing half-edges that belong to a ruined agent the *ruinous* half-edges.

We consider the following loss reveal process, extending the risk processes of Section 4.2.1.

In this process, losses coming from ruined agents are revealed one by one. At each loss reveal we look at the interaction between an outgoing half-edge of a ruined agent (ruinous outgoing half-edge) with an incoming half-edge of its counterparty. By virtue of the configuration model, this is chosen uniformly at random among all remaining incoming half-edges. When the ruinous outgoing half-edge is matched, the counterparty incurs a random loss, drawn from a distribution depending on the characteristics class $x \in \mathcal{X}$ of the counterparty. If this amount of loss is larger than the remaining capital, this agent will become ruined and all its outgoing half-edges become ruinous. Note that the loss reveal process stops when all ruinous outgoing half-edges have been matched. We use the notation $W_n(t)$ to denote the number of remaining (unrevealed) ruinous outgoing half-edges at time t. The contagion thus stops at the first time when $W_n(t) = 0$.

We consider a general loss reveal intensity process, denoted by $\mathcal{R}_n(t)$, to describe the intensity of loss reveal at time t. Namely, if a loss is revealed at time $t_1 \in \mathbb{R}_+$, then we wait an exponential time with parameter $\mathcal{R}_n(t_1)$ for the next loss reveal. Note that $\mathcal{R}_n(t)$ could depend on the state of risk processes at time t. In particular, the networked risk processes of Section 4.2.1 in configuration model would be equivalent to this loss reveal process by setting $\mathcal{R}_n(t) = \beta W_n(t)$. Indeed, each counterparty of a ruinous half-edge will be revealed after an exponential time with parameter β. When there are $W_n(t)$ such unrevealed counterparties, the next reveal will be given by the minimum of these exponential times.

Our model allows for a general form of the loss reveal intensity $\mathcal{R}_n(t)$, provided that the following condition holds. Let $\tau_n^\star$ denote the stopping time defined as the first time when the above loss reveal process ends. This is the first time such that $W_n(\tau_n^\star) = 0$.

Assumption 4.3. *We assume that the loss intensity function $\mathcal{R}_n$ satisfies $\mathcal{R}_n(t) = 0$ for $t > \tau_n^\star$, and $\mathcal{R}_n(t) = n\mathfrak{R}(t) + o_p(n)$ for $t \leqslant \tau_n^\star$ with $\mathfrak{R}(t)$ continuous, positive and $o_p(n)$ is uniform for $t \leqslant \tau_n^\star$.*

By Theorem 4.5, this assumption holds for the risk processes of Section 4.2.1. In the next section we state the limit theorems for the general loss reveal process satisfying Assumption 4.3 and then we apply them to the particular risk processes of Section 4.2.1.

4.3 Main Theorems

4.3.1 Asymptotic analysis of the general loss reveal process

We consider the loss reveal process of Section 4.2.3 satisfying Assumption 4.3 on the random graph $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ which satisfies Assumptions 3.3-4.2. Let us denote by $\mathcal{I}_n(t)$ a Cox process with intensity $\mathcal{R}_n(t)$ at time t. This represents the total number of ruinous outgoing half-edges revealed before time t. Then $\mathcal{I}_n(\infty)$ represents the total number of ruinous outgoing half-edges that will be revealed if the reveal process continues forever. Since the total number of reveals is bounded from above by the total number of links in the network, we need to stop the Cox process at $\mathcal{I}_n^\star = \mathcal{I}_n(\infty) \wedge (n\lambda^{(n)})$, where for $a, b \in \mathbb{R}$, we denote by $a \wedge b := \min\{a, b\}$. We define

$$t_{\mathfrak{R}}(\lambda) := \inf\{t \geqslant 0 : \int_0^t \mathfrak{R}(s)ds \geqslant \lambda\}.$$

By convention, if $\|\mathfrak{R}\|_{L^1} \leqslant \lambda$ we set $t_{\mathfrak{R}}(\lambda) := \infty$.

For $z \in [0,1]$, denote by $T_n(z)$ the time needed to reveal $\lceil nz \rceil$ ruinous outgoing half-edges. The following lemma gives the asymptotic results on $\mathcal{I}_n(t)$, $\mathcal{I}_n^\star$ and $T_n(z)$.

Lemma 4.2. *Under Assumptions 4.1-4.2, and for any given loss intensity function $\mathcal{R}_n$ satisfying Assumption 4.3, we have as $n \to \infty$,*

$$\sup_{t \geqslant 0} \Big| \frac{\mathcal{I}_n(t)}{n} - \int_0^{t \wedge t_{\mathfrak{R}}(\lambda)} \mathfrak{R}(s)ds \Big| \xrightarrow{p} 0, \quad \text{and} \quad \frac{\mathcal{I}_n^\star}{n} \xrightarrow{p} \|\mathfrak{R}\|_{L^1} \wedge \lambda. \tag{4.5}$$

Further, for all $0 \leqslant a < b < \|\mathfrak{R}\|_{L^1} \wedge \lambda$ and as $n \to \infty$,

$$T_n(b) - T_n(a) \xrightarrow{p} t_{\mathfrak{R}}(b) - t_{\mathfrak{R}}(a). \tag{4.6}$$

The proof of lemma is provided in Section 4.4.2.

For each $x \in \mathcal{X}$, we denote by $\boldsymbol{L}_x := (L_x^{(1)}, \dots, L_x^{(d_x^+)})$ the sequence of independent random losses with distribution F_x and let $\boldsymbol{\ell}_x = (\ell_x^{(1)}, \ell_x^{(2)}, \dots, \ell_x^{(d_x^+)})$ be a realization of $\boldsymbol{L}_x$. For a given $\boldsymbol{\ell}_x$ and a given initial shock ϵ_x, define for all $\theta = 0, 1, \ldots, d_x^+$,

$$\tau_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x) := \inf\{t \geqslant 0 : \gamma_x(1-\epsilon_x) + \alpha_x(t) - \delta_x \geqslant \sum_{i=1}^{\theta} \ell_x^{(i)}\}. \tag{4.7}$$

The function τ can be seen as the time threshold function of the loss $\boldsymbol{\ell}_x$. Indeed, for initial shock ϵ and upcoming sequence of losses $\boldsymbol{\ell}_x$, the threshold function $\tau_{x,k}(\epsilon, \boldsymbol{\ell}_x)$ is the smallest time needed for a firm of type x to have enough capital for absorbing these incoming k losses. Recall that a firm of type x has a capital growth function α_x and external debt δ_x. Note that $\tau_{x,0}(\epsilon_x, \boldsymbol{\ell}_x) > 0$ denotes the event that agent of type x initially becomes ruined under the shock ϵ_x.

Example 4.3. *Consider the simple case where there is no recovery for agents, i.e. $\alpha_x = 0$ for all $x \in \mathcal{X}$, and the amount of loss for each agent in the same class $x \in \mathcal{X}$ is constant ℓ_x. Then by the definition of $\tau_{x,\theta}$, we have*

$$\tau_{x,\theta} = \begin{cases} 0 & \text{if} \quad \theta < \lceil \frac{\gamma_x(1-\epsilon_x)-\delta_x}{\ell_x} \rceil, \\ \infty & \text{if} \quad \theta \geqslant \lceil \frac{\gamma_x(1-\epsilon_x)-\delta_x}{\ell_x} \rceil. \end{cases}$$

This would be equivalent to type-dependent threshold contagion model which extends the bootstrap percolation model. In bootstrap percolation model, the threshold is fixed for all nodes, i.e. $\lceil \gamma_x/\ell_x \rceil = \theta$ for all $x \in \mathcal{X}$. We refer to [8, 10, 44] for results on bootstrap percolation in configuration model.

For a given positive density function $\mathfrak{R} : \mathbb{R}_0^+ \to \mathbb{R}^+$ with $\|\mathfrak{R}\|_{L^1} < \infty$, $x \in \mathcal{X}$ and $\theta = 0, 1, \ldots, d_x^+$, we let the survival probability be (for all $t \geqslant 0$)

$$P_{x,\theta}^{\mathfrak{R}}(t, \epsilon_x, \boldsymbol{\ell}_x) := \mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{\ell}_x) = 0, U_{(1)}^{\mathfrak{R},t} > \tau_{x,1}(\epsilon_x, \boldsymbol{\ell}_x), \ldots, U_{(\theta)}^{\mathfrak{R},t} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x)), \tag{4.8}$$

with the convention $P_{x,0}^{\mathfrak{R}}(t, \epsilon_x, \boldsymbol{\ell}_x) := \mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{\ell}_x) = 0)$ for all $x \in \mathcal{X}$, where $U_{(1)}^{\mathfrak{R},t}, U_{(2)}^{\mathfrak{R},t}, \ldots, U_{(\theta)}^{\mathfrak{R},t}$ are the order statistics of θ i.i.d. random variables $\{U_i^{\mathfrak{R},t}\}_{i=1,\ldots,\theta}$ with distribution

$$\mathbb{P}(U_i^{\mathfrak{R},t} \leqslant y) = \frac{\int_0^y \mathfrak{R}(s)ds}{\int_0^t \mathfrak{R}(s)ds}, \qquad y \leqslant t. \tag{4.9}$$

Knowing that a loss arrives before time t, the probability that it arrives before time y is given by (4.9). Then for a fixed sequence of given losses and initial shock, Equation 4.8 represents the probability that the firm survives at time t, given that there are θ by that time (It must then be that the ordered losses arrived after the successive threshold times).

Remark 4.4. *The joint probability density of the order statistics $U_{(1)}^{\mathfrak{R},t}, U_{(2)}^{\mathfrak{R},t}, \ldots, U_{(\theta)}^{\mathfrak{R},t}$ is given by*

$$f_{U_{(1)}^{\mathfrak{R},t}, U_{(2)}^{\mathfrak{R},t}, \ldots, U_{(\theta)}^{\mathfrak{R},t}}(u_1, u_2, \ldots, u_\theta) = \theta!(\int_0^t \mathfrak{R}(s)ds)^{-\theta} \prod_{i=1}^{\theta} \mathfrak{R}(u_i),$$

for all $u_1, \ldots, u_\theta \in [0, t]$.

By integrating the conditional survival probability in Equation (4.8) with respect to the probability density function of ϵ_x and $\boldsymbol{\ell}_x$, we obtain the survival probability at time t for any agent of type x with θ incoming losses absorbed by t, denoted by $\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)$. This is defined by $\mathcal{S}^{\mathfrak{R}}_{x,0}(t) := 1 - q_{x,0}$ and

$$\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t) := \int P^{\mathfrak{R}}_{x,\theta}(t, \epsilon_x, \boldsymbol{\ell}_x) f^{(\epsilon)}_x(\epsilon_x) f_x(\ell^{(1)}_x) f_x(\ell^{(2)}_x) \cdots f_x(\ell^{(\theta)}_x) d\epsilon_x d\ell^{(1)}_x d\ell^{(2)}_x \cdots d\ell^{(\theta)}_x,$$

for $\theta = 1, \ldots, d^+_x$. It can be equally written as

$$\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t) = P^{\mathfrak{R}}_{x,\theta}(t, \epsilon_x, \boldsymbol{L}_x)] = \mathbb{E}\big[\mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{L}_x) = 0, U^{\mathfrak{R},t}_{(1)} > \tau_{x,1}(\epsilon_x, \boldsymbol{L}_x), \ldots, U^{\mathfrak{R},t}_{(\theta)} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{L}_x)), \quad (4.10)$$

where $\boldsymbol{L}_x := (L^{(1)}_x, \ldots, L^{(d^+_x)}_x)$ is a sequence of independent random losses with distribution F_x and ϵ_x is an independent random variable with distribution $F^{(\epsilon)}_x$.

For a given positive function $\mathfrak{R} : \mathbb{R}^+_0 \to \mathbb{R}^+$, we define

$$\phi^{\mathfrak{R}}(t) := \frac{\int_0^{t \wedge t_{\mathfrak{R}}(\lambda)} \mathfrak{R}(s) ds}{\lambda}, \tag{4.11}$$

so that the binomial probability $b(d^+_x, \phi^{\mathfrak{R}}(t), \theta)$ represents the probability an agent of type x suffers θ losses prior to time t.

For $t \geqslant 0$ and given $\mathfrak{R}$, we define the following functions which will be shown to be the limiting fractions of surviving and defaulted agents, respectively:

$$f^{\mathfrak{R}}_S(t) := \sum_{x \in \mathcal{X}} \mu_x \sum_{\theta=0}^{d^+_x} b(d^+_x, \phi^{\mathfrak{R}}(t), \theta) \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t), \quad f^{\mathfrak{R}}_D(t) = 1 - f^{\mathfrak{R}}_S(t). \tag{4.12}$$

We further define the following function which will be shown to be fraction of remaining (unrevealed) ruinous outgoing half-edges at time t (characterizing the contagion stopping time):

$$f^{\mathfrak{R}}_W(t) := \lambda(1 - \phi^{\mathfrak{R}}(t)) - \sum_{x \in \mathcal{X}} \mu_x d^-_x \sum_{\theta=0}^{d^+_x} b(d^+_x, \phi^{\mathfrak{R}}(t), \theta) \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t). \tag{4.13}$$

Now we can state the following limit theorem regarding the fraction of solvent and defaulted agents at time t. Recall that $\tau^\star_n$ is the stopping time at which the ruin propagation stops, i.e., this is the first time such that $W_n(\tau^\star_n) = 0$.

Theorem 4.5. *Under Assumptions 4.1-4.2, and for any given loss intensity function $\mathcal{R}_n$ satisfying Assumption 4.3, we have as $n \to \infty$,*

$$\sup_{t \leqslant \tau^\star_n} \Big| \frac{S^{(n)}_{x,\theta}(t)}{n} - \mu_x b(d^+_x, \phi^{\mathfrak{R}}(t), \theta) \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t) \Big| \xrightarrow{p} 0.$$

Further, as $n \to \infty$,

$$\sup_{t \leqslant \tau_n^\star} \Big| \frac{S^{(n)}(t)}{n} - f_S^{\mathfrak{R}}(t) \Big| \xrightarrow{p} 0, \quad \sup_{t \leqslant \tau_n^\star} \Big| \frac{D^{(n)}(t)}{n} - f_D^{\mathfrak{R}}(t) \Big| \xrightarrow{p} 0,$$

and the process W_n satisfies

$$\sup_{t \leqslant \tau_n^\star} \Big| \frac{W_n(t)}{n} - f_W^{\mathfrak{R}}(t) \Big| \xrightarrow{p} 0.$$

The proof of theorem is provided in Section 4.4.3.

Remark 4.6. *The above theorem generalizes the results of [29] to the case of a general loss reveal intensity function $\mathcal{R}_n(t)$ satisfying Assumption 4.3. In contrast, in [29], the reveal intensity is assumed to be constant equal to the size of network n and the recovery rate is proportional to the agent's connectivity d_x^+, for each type $x \in \mathcal{X}$ (the type is simply the degrees of each agent in [29]). Namely, $\mathcal{R}_n(t) = n$ and $\alpha_x(t) = \alpha d_x^+/\lambda t$ for all $x \in \mathcal{X}$. In this case, $\mathfrak{R}(t) = 1$ for $t \in [0, \lambda]$ and $\mathfrak{R}(t) = 0$ for $t > \lambda$. Thus $\phi^{\mathfrak{R}}(t) = t/\lambda$ and for $\theta = 1, \dots, d_x^+$, $U_{(1)}^{\mathfrak{R},t}, U_{(2)}^{\mathfrak{R},t}, \dots, U_{(\theta)}^{\mathfrak{R},t}$ become the order statistics of θ i.i.d. uniform random variables on $[0, \lambda]$. Also, the time threshold $\tau_{x,k}$ does not have a dependence on the loss $\boldsymbol{\ell}_x$ as in our definition (4.7). It simplifies to*

$$\tau_{x,\theta} := \inf\{t \geqslant 0 : \Theta_x + \alpha t d_x^+/\lambda \geqslant \theta\},$$

where Θ_x is the random initial default threshold $\mathbb{P}(\Theta_x = k) = q_{x,k}$ as defined in Remark 4.1. In this case, we can regard the default threshold as the number of losses each agent could absorb and we could recover the results of [29].

In the case when there is no growth in the network and $\alpha_x(t) = 0$, it is more convenient to characterize the above limit functions through the threshold distribution functions $q_{x,\theta}$ and $\bar{q}_x$ which could be defined as (similar to Remark 4.1)

$$q_{x,\theta}(t) := \mathbb{P}\big(L_x^{(1)} + \cdots + L_x^{(\theta-1)} < \gamma_x(1-\epsilon_x) + \alpha_x(t) - \delta_x \leqslant L_x^{(1)} + \cdots + L_x^{(\theta)}\big),$$

for $\theta = 1, \dots, d_x^+$, and

$$\bar{q}_x := \mathbb{P}\big(L_x^{(1)} + \cdots + L_x^{(d_x^+)} < \gamma_x(1-\epsilon_x) + \alpha_x(t) - d_x\big) = 1 - \sum_{\theta=0}^{d_x^+} q_{x,\theta}.$$

For $t \geqslant 0$, we define the functions (which are the simplified versions of those in (4.12) and respectively (4.13))

$$\widehat{f}_S^{\mathfrak{R}}(t) := \sum_{x \in \mathcal{X}} \mu_x \Big[\sum_{\theta=1}^{d_x^+} q_{x,\theta} \beta\big(d_x^+, 1 - \phi^{\mathfrak{R}}(t), d_x^+ - \theta + 1\big) + \bar{q}_x \Big], \quad \widehat{f}_D^{\mathfrak{R}}(t) = 1 - \widehat{f}_S^{\mathfrak{R}}(t),$$

$$\widehat{f}_W^{\mathfrak{R}}(t) := \lambda(1 - \phi^{\mathfrak{R}}(t)) - \sum_{x \in \mathcal{X}} \mu_x d_x^- \Big[\sum_{\theta=1}^{d_x^+} q_{x,\theta} \beta\big(d_x^+, 1 - \phi^{\mathfrak{R}}(t), d_x^+ - \theta + 1\big) + \bar{q}_x \Big].$$

As a corollary of Theorem 4.5, we have the following theorem.

Theorem 4.7. *Suppose there is no growth in the network, i.e., $\alpha_x = 0$ for all $x \in \mathcal{X}$. Under Assumptions 4.1-4.2, and for any given loss intensity function $\mathcal{R}_n$ satisfying Assumption 4.3, we have as $n \to \infty$,*

$$\sup_{t \leqslant \tau_n} \Big| \frac{S_{x,\theta}^{(n)}(t)}{n} - \mu_x b\big(d_x^+, \phi^{\mathfrak{R}}(t), \theta\big) \Big(\sum_{\delta=\theta+1}^{d_x^+} q_{x,\delta} + \bar{q}_x \Big) \Big| \xrightarrow{p} 0$$

for all $x \in \mathcal{X}$ and $\theta = 1, \ldots, d_x^+$. Further, as $n \to \infty$,

$$\sup_{t \leqslant \tau_n} \Big| \frac{S^{(n)}(t)}{n} - \widehat{f}_S^{\mathfrak{R}}(t) \Big| \xrightarrow{p} 0, \quad \sup_{t \leqslant \tau_n} \Big| \frac{D^{(n)}(t)}{n} - \widehat{f}_D^{\mathfrak{R}}(t) \Big| \xrightarrow{p} 0,$$

and the process W_n satisfies

$$\sup_{t \leqslant \tau_n} \Big| \frac{W_n(t)}{n} - \widehat{f}_W^{\mathfrak{R}}(t) \Big| \xrightarrow{p} 0.$$

The proof of above theorem is provided in Section 4.4.4.

Theorem 4.7 generalizes the limit theorem of [20] to the dynamic case. It also generalizes the law of large numbers result in Chapter 2, where the authors consider default cascades in configuration model such that all outgoing half-edges are assigned with an exponential clock with parameter one, leading to a loss reveal intensity function equal to the number of remaining outgoing half-edges at time $t \geqslant 0$ and satisfying $\mathcal{R}_n(t) = n\lambda e^{-t} + o_p(n)$.

In order to determine the ruin probabilities, we need to study the stopping time $\tau_n^\star$ when there are no more ruinous outgoing links in the system. Since from Theorem 4.5, the fraction of remaining ruinous outgoing half-edges converges to $f_W^{\mathfrak{R}}(t)$, we define

$$t_{\mathfrak{R}}^\star := \inf\big\{t \in [0,1] : f_W^{\mathfrak{R}}(t) = 0\big\}. \tag{4.14}$$

We say that $t_{\mathfrak{R}}^\star < \infty$ is a stable solution of $f_W^{\mathfrak{R}}(t) = 0$ if there exists a small $\epsilon > 0$ such that $f_W^{\mathfrak{R}}(t)$ is negative on $[t_{\mathfrak{R}}^\star, t_{\mathfrak{R}}^\star + \epsilon)$. We have the following lemma.

Lemma 4.8. *Under Assumptions 4.1-4.2, and for any given loss intensity function $\mathcal{R}_n$ satisfying Assumption 4.3, we have as $n \to \infty$:*

- *If $t_{\mathfrak{R}}^\star < \infty$ is a stable solution of $f_W^{\mathfrak{R}}(t) = 0$, then $\tau_n^\star \xrightarrow{p} t_{\mathfrak{R}}^\star$.*
- *If $t_{\mathfrak{R}}^\star = \infty$, then $\tau_n^\star \xrightarrow{p} \infty$.*

The proof of lemma is provided in Section 4.4.5.

We are now ready to provide the limit theorem about the final ruin probabilities. As a corollary of Theorem 4.5 and Lemma 4.8 the following holds.

Theorem 4.9. *Under Assumptions 4.1-4.2, and for any given loss intensity function $\mathcal{R}_n$ satisfying Assumption 4.3, we have as $n \to \infty$:*

(i) If $\int_0^{t^\star_{\mathfrak{R}}} \mathfrak{R}(s)ds = \lambda$, *then asymptotically all agents are ruined by the end of the loss propagation process, i.e.*

$$D^{(n)}(\tau_n^\star) = n - o_p(n).$$

(ii) If $t^\star_{\mathfrak{R}} < \infty$ *is a stable solution of* $f_W^{\mathfrak{R}}(t) = 0$ *and* $\int_0^{t^\star_{\mathfrak{R}}} \mathfrak{R}(s)ds < \lambda$, *then the ruin probability of an agent of type* $x \in \mathcal{X}$ *converges to*

$$\frac{D_x^{(n)}(\tau_n^\star)}{n\mu_x^{(n)}} \xrightarrow{p} 1 - \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t^\star_{\mathfrak{R}}),$$

and the total number of ruined agents satisfies

$$D^{(n)}(\tau_n^\star) = n\sum_{x\in\mathcal{X}} \mu_x(1 - \sum_{\theta=0}^{d_x^+} \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t^\star_{\mathfrak{R}})) + o_p(n).$$

(iii) If $t^\star_{\mathfrak{R}} = \infty$ *and* $\|\mathfrak{R}\|_{L^1} < \lambda$, *then the ruin probability of an agent of type* $x \in \mathcal{X}$ *converges to*

$$\frac{D_x^{(n)}(\tau_n^\star)}{n\mu_x^{(n)}} \xrightarrow{p} 1 - \sum_{\theta=0}^{d_x^+} b\big(d_x^+, \|\mathfrak{R}\|_{L^1}/\lambda, \theta\big)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(\infty),$$

and the total number of ruined agents satisfies

$$D^{(n)}(\tau_n^\star) = n\sum_{x\in\mathcal{X}} \mu_x(1 - \sum_{\theta=0}^{d_x^+} b\big(d_x^+, \|\mathfrak{R}\|_{L^1}/\lambda, \theta\big)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(\infty)) + o_p(n),$$

where $\mathcal{S}^{\mathfrak{R}}_{x,\theta}(\infty)$ *denotes the limit of* $\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)$ *as* $t \to \infty$.

The proof of theorem is provided in Section 4.4.6.

4.3.2 Ruin probabilities for the networked risk processes

We consider the networked risk processes of Section 4.2.1 on the random graph $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$. In this case the loss reveal intensity function is totally determined by the number of ruinous outgoing half-edges, namely $\mathcal{R}_n(t) = \beta W_n(t)$.

In the previous section we have assumed that the loss intensity function $\mathcal{R}_n$ has a limit function $\mathfrak{R}$ which satisfies Assumption 4.3. We will now show that for the networked risk processes of Section 4.2.1, Assumption 4.3 holds and there exists a unique limit function $\mathfrak{R}^\star$. We will take advantage of Theorem 4.9 to show that $\mathfrak{R}^\star$ can be characterized as a fixed point solution, representing the limit of remaining ruinous links.

To obtain this existence and uniqueness result for $\mathfrak{R}^\star$, we need to consider a second moment condition for the degrees of the random graph $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$.

Assumption 4.4. *We assume that, as $n \to \infty$, $\sum_{i\in[n]}(d_i^+ + d_i^-)^2 = O(n)$.*

In particular, the above assumption implies (by uniform integrability) Assmption 4.2 and $\lambda^{(n)} \to \lambda$ as $n \to \infty$. In this case, since $\liminf_{n\longrightarrow\infty} \mathbb{P}(\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$ simple$) > 0$, our limit theorems could be transferred to the uniformly distributed random graph with these degree sequences $\mathcal{G}_*^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$, see e.g., [194].

We have the following theorem.

Theorem 4.10. *Let $\mathbb{L}_\lambda(\mathbb{R}^+)$ be the space of all continuous positive integrable functions f with $\|f\|_1 \leqslant \lambda$. Suppose that the loss reveal intensity satisfies $\mathcal{R}_n(t) = \beta W_n(t)$ and the network sequence $\{\mathcal{G}_n\}_{n\in\mathbb{N}}$ satisfies Assumptions 4.1 and 4.4. Then we have:*

(i) There exists a unique solution $\mathfrak{R}^\star$ in $\mathbb{L}_\lambda(\mathbb{R}^+)$ with an initial value $\mathfrak{R}^\star(0) = \beta\sum_{x\in\mathcal{X}}\mu_x d_x^-(1-q_{x,0})$ to the fixed point equation $\mathfrak{R} = \beta\Psi(\mathfrak{R})$, where $\Psi : \mathbb{L}_\lambda(\mathbb{R}^+) \to \mathbb{L}_\lambda(\mathbb{R}^+)$ is the map

$$\Psi(\mathfrak{R})(t) = \lambda(1-\phi^{\mathfrak{R}}(t)) - \sum_{x\in\mathcal{X}}\mu_x d_x^- \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t).$$

(ii) As $n \to \infty$, we have

$$\sup_{t\leqslant\tau_n^\star}\Big|\frac{\beta W_n(t)}{n} - \mathfrak{R}^\star(t)\Big| \xrightarrow{p} 0,$$

and consequently,

$$\sup_{t\leqslant\tau_n^\star}\Big|\frac{S^{(n)}(t)}{n} - f_S^{\mathfrak{R}^\star}(t)\Big| \xrightarrow{p} 0 \quad \text{and} \quad \sup_{t\leqslant\tau_n^\star}\Big|\frac{D^{(n)}(t)}{n} - f_D^{\mathfrak{R}^\star}(t)\Big| \xrightarrow{p} 0.$$

The proof of the theorem is provided in Section 4.4.7.

We also have the following lemma which guarantees that $\tau_n^\star \xrightarrow{p} \infty$ in this model.

Lemma 4.11. *Let all assumptions in Theorem 4.10 hold and $\mathfrak{R}^\star(0) = \beta\sum_{x\in\mathcal{X}}\mu_x d_x^-(1-q_{x,0}) > 0$. Then we have $\tau_n^\star > (1-\epsilon)\log n$ with high probability for any $\epsilon > 0$.*

As a direct consequence of Theorem 4.9 and Proposition 4.11, we have $\tau_n^\star \xrightarrow{p} \infty$ as $n \to \infty$, and thus the final state belongs to point (ii) in Theorem 4.10. In this case, $\mathcal{R}_n(t) = \beta W_n(t) = n\mathfrak{R}^\star + o_p(n)$. If $\|\mathfrak{R}^\star\|_{L^1} = \lambda$, then asymptotically all agents become ruined during the cascade. Otherwise, if $\|\mathfrak{R}^\star\|_{L^1} < \lambda$, for any type $x \in \mathcal{X}$, the ruin probability of agent of type x is

$$\frac{D_{x,\theta}^{(n)}(\tau_n^\star)}{n\mu_x^{(n)}} = 1 - \sum_{\theta=0}^{d_x^+} b(d_x^+, \|\mathfrak{R}^\star\|_{L^1}/\lambda, \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}^\star}(\infty) + o_p(1),$$

and the total number of ruined agents satisfies

$$D^{(n)}(\tau_n^\star) = n\sum_{x\in\mathcal{X}}\mu_x(1 - \sum_{\theta=0}^{d_x^+} b(d_x^+, \|\mathfrak{R}^\star\|_{L^1}/\lambda, \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}^\star}(\infty)) + o_p(n).$$

4.4 Proofs

Before proving our main theorems, we will introduce some preliminary definitions and lemmas.

4.4.1 Some preliminary results

We begin by asserting the following claim. The proof is straightforward and we omit it.

Claim 4.12. *Let* $\boldsymbol{U}_x^{\mathfrak{R}} = (U_{x,1}^{\mathfrak{R}}, \dots, U_{x,d_x^+}^{\mathfrak{R}})$ *be a vector of* d_x^+ *i.i.d. random variables with common distribution (similiar to Equation 4.9 replacing* t *by* $t_{\mathfrak{R}}(\lambda)$*)*

$$\mathbb{P}(U_{x,i}^{\mathfrak{R}} \leqslant y) = \frac{\int_0^y \mathfrak{R}(s)ds}{\int_0^{t_{\mathfrak{R}}(\lambda)} \mathfrak{R}(s)ds}, \ y \leqslant t_{\mathfrak{R}}(\lambda), i = 1, \dots, d_x^+,$$

for all $x \in \mathcal{X}$*. Then* $U_{(1)}^{\mathfrak{R},t}, U_{(2)}^{\mathfrak{R},t}, \dots, U_{(k)}^{\mathfrak{R},t}$ *have the same distribution as that of the first* k *order statistics of* $\boldsymbol{U}_x^{\mathfrak{R}}$ *conditioned on* $t \in [U_{x,(k)}^{\mathfrak{R}}, U_{x,(k+1)}^{\mathfrak{R}})$*, where* $U_{x,(1)}^{\mathfrak{R}} \leqslant U_{x,(2)}^{\mathfrak{R}} \leqslant \dots U_{x,(d_x^+)}^{\mathfrak{R}}$ *are the order statistics of elements of* $\boldsymbol{U}_x^{\mathfrak{R}}$*. Thus the probability measure of* $U_{(k)}^{\mathfrak{R},t}$*, for all* $k = 1, \dots, d_x^+$ *and* $t \geqslant 0$*, can be generated by the vector* $\boldsymbol{U}_x^{\mathfrak{R}}$*.*

Let $\boldsymbol{L}_x$ be a vector of d_x^+ i.i.d. random losses to an agent of type x and $\boldsymbol{L}_{x,k}$ be the subvector of first k positions. From (4.10), $\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t)$ can be regarded as $\mathbb{P}_x H_{x,\theta,t}$, where $H_{x,\theta,t} : [0, +\infty) \times (\mathbb{R}^+)^{d_x^+} \times [0, +\infty)^{d_x^+} \mapsto \mathbb{R}$ is a measurable function defined as

$$H_{x,\theta,t}(\epsilon_x, \boldsymbol{\ell}_x, \boldsymbol{u}_x) := \mathbb{1}\{u_{(k)} \leqslant t < u_{(k+1)}\}\mathbb{1}\{\tau_{x,0}(\epsilon_x, \boldsymbol{\ell}_x) = 0, u_{(1)} > \tau_{x,1}(\epsilon_x, \boldsymbol{\ell}_x), \dots, u_{(\theta)} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x)\},$$

where $u_{(1)}, u_{(2)}, \dots, u_{(d_x^+)}$ is the order statistics of $\boldsymbol{u}_x$ and $\mathbb{P}_x$ is the probability measure on $[0, +\infty) \times (\mathbb{R}^+)^{d_x^+} \times [0, +\infty)^{d_x^+}$ generated by $(\epsilon_x, \boldsymbol{L}_x, \boldsymbol{U}_x^{\mathfrak{R}})$.

Let us define the class of functions $\mathcal{H}_{x,\theta}$ as the collection of $H_{x,\theta,t}$ for all $t \in \mathbb{R}$, i.e., $\mathcal{H}_{x,\theta} := \{H_{x,\theta,t}, t \geqslant 0\}$. We will show that for each $x \in \mathcal{X}$ and $\theta \leqslant d_x^+$, the class $\mathcal{H}_{x,\theta}$ is a Glivenko-Cantelli class with respect to $\mathbb{P}_x$ defined as follows.

Definition 4.13. *A class of functions (or sets)* $\mathcal{H}$ *is called a strong Glivenko-Cantelli class with respect to a probability measure* $\mathbb{P}$ *if*

$$\sup_{H \in \mathcal{H}} |\mathbb{P}_n H - \mathbb{P}H| \xrightarrow{a.s.} 0, \qquad as \ \ n \to \infty,$$

where $\mathbb{P}_n$ *is the empirical measure of* $\mathbb{P}$*.*

We also need to introduce the following definitions.

For a collection of subsets of Ω, denoted by $\mathcal{H}$, with Ω being a space (usually a sample space), the n-th shattering coefficient of $\mathcal{H}$ is defined by

$$S_{\mathcal{H}}(n) := \max_{x_1, \dots, x_n \in \Omega} \operatorname{card}\{\{x_1, \dots, x_n\} \cap \mathcal{A}, \mathcal{A} \in \mathcal{H}\},$$

where card{.} denotes the cardinality of the set. Then the Vapnik–Chervonenkis dimension (or VC dimension) of $\mathcal{H}$ is defined as

$$\mathrm{VC}(\mathcal{H}) := \max\{n \geqslant 1 : S_{\mathcal{H}}(n) = 2^n\}.$$

We need the following Glivenko-Cantelli theorem for finite VC classes; see e.g. [109] for more details.

Theorem 4.14 (Glivenko-Cantelli theorem for finite VC classes)**.** *Let $\mathbb{P}$ be a probability measure on a Polish space Ω and $\mathcal{H}$ be a collection of subsets of Ω with $\mathrm{VC}(\mathcal{H}) < \infty$. Then $\mathcal{H}$ is a strong Glivenko-Cantelli class with respect to $\mathbb{P}$.*

As a result of above theorem, we have the following lemma.

Lemma 4.15. *For each $x \in \mathcal{X}$ and $\theta = 0, 1, \ldots, d_x^+$, $\mathcal{H}_{x,\theta}$ is a strong Glivenko-Cantelli class with respect to $\mathbb{P}_x$.*

Proof. For each $x \in \mathcal{X}$ and $\theta = 0, 1, \ldots, d_x^+$, it can be deduced from Claim 4.12 that $\mathbb{P}_x$ is a probability measure generated by $(\epsilon_x, \boldsymbol{L}_x, \boldsymbol{U}_x^{\mathfrak{R}})$, and is defined on a Polish space. It can be verified that $\mathrm{VC}(\mathcal{H}_{x,\theta}) = 2$, making it a finite VC class. The lemma follows then from Theorem 4.14. ☐

We will also use the following well known result, see e.g. [186] for a proof, which offers us a way to convert the loss reveal process into a death process with an initial number $\mathcal{I}_n^\star$ and i.i.d. lifetimes.

Lemma 4.16. *Let $\{N(t) : t \geqslant 0\}$ be an inhomogeneous Poisson process with an intensity function $\mathfrak{R}(t)$ and arrival times $\{\sigma_k, k \geqslant 1\}$. Given any fixed $t > 0$ and conditional on m arrivals before time t, the random vector $(\sigma_1, \sigma_2, \ldots, \sigma_m)$ has the same distribution as the random vector $(Y_{(1)}, Y_{(2)}, \ldots, Y_{(m)})$, where $Y_{(i)}$ is the i-th order statistic of m i.i.d. random variables with probability density function $\mathfrak{R}(s)\left(\int_0^t \mathfrak{R}(u)du\right)^{-1}$.*

We are now ready to present the proofs of our main theorems.

4.4.2 Proof of Lemma 4.2

We first prove that for any $t \geqslant 0$, as $n \to \infty$,

$$\frac{\mathcal{I}_n(t)}{n} \xrightarrow{p} \lambda\phi^{\mathfrak{R}}(t).$$

Let Δs be a small time interval and let $m = \lceil t/\Delta s \rceil$ represent the number of time intervals between $[0, t]$. Notice that, as $n \to \infty$,

$$\frac{\mathbb{E}[\mathcal{I}_n(t)]}{n} = \frac{1}{n}\sum_{i=1}^{m} \mathbb{E}[\mathcal{I}_n(i\Delta s) - \mathcal{I}_n((i-1)\Delta s)] = \sum_{i=1}^{m} \Delta s(\mathfrak{R}(i\Delta s) + C\delta) + o(1),$$

where $\delta := \sup_{u,v\in[0,t],|u-v|\leqslant\Delta s} |\mathfrak{R}(u) - \mathfrak{R}(v)|$, is a small error and C is some constant. Hence for $t \leqslant t_{\mathfrak{R}}(\lambda)$ as $\Delta s \to 0$, by the continuity of $\mathfrak{R}$, we have

$$\frac{\mathbb{E}[\mathcal{I}_n(t)]}{n} \to \int_0^t \mathfrak{R}(s)ds.$$

On the other hand, by the law of total variance, we have

$$\begin{aligned}
&\mathrm{Var}[\mathcal{I}_n(i\Delta s) - \mathcal{I}_n((i-1)\Delta s)] \\
&\quad = \mathrm{Var}\big[\mathbb{E}[\mathcal{I}_n(i\Delta s) - \mathcal{I}_n((i-1)\Delta s)|\mathcal{R}_n]\big] + \mathbb{E}\big[\mathrm{Var}[\mathcal{I}_n(i\Delta s) - \mathcal{I}_n((i-1)\Delta s)|\mathcal{R}_n]\big] \\
&\quad = \mathrm{Var}\Big[\int_{(i-1)\Delta s}^{i\Delta s} \mathcal{R}_n(u)du\Big] + \mathbb{E}\Big[\int_{(i-1)\Delta s}^{i\Delta s} \mathcal{R}_n(u)du\Big].
\end{aligned}$$

It therefore follows that for $t \leqslant t_{\mathfrak{R}}(\lambda)$,

$$\begin{aligned}
\mathrm{Var}\Big(\frac{\mathcal{I}_n(t)}{n}\Big) &= \frac{1}{n^2}\Big(\mathrm{Var}\Big[\int_0^t \mathcal{R}_n(u)du\Big] + \mathbb{E}\Big[\int_0^t \mathcal{R}_n(u)du\Big]\Big) \\
&= \frac{1}{n}\int_0^t \mathfrak{R}(u)du + o(1) = o(1).
\end{aligned}$$

Recall the definition of $\phi^{\mathfrak{R}}$,

$$\phi^{\mathfrak{R}}(t) := \frac{\int_0^{t\wedge t_{\mathfrak{R}}(\lambda)} \mathfrak{R}(s)ds}{\lambda},$$

Then by Chebysev's inequality we have, as $n \to \infty$, for any $t \geqslant 0$,

$$\frac{\mathcal{I}_n(t)}{n} \xrightarrow{p} \lambda\phi^{\mathfrak{R}}(t). \tag{4.15}$$

As a consequence, by letting $t \to \infty$, the final fraction of revealed ruinous outgoing half-edges converges to $\|\mathfrak{R}\|_{L^1}$ in probability, i.e.,

$$\frac{\mathcal{I}_n^\star}{n} \xrightarrow{p} \lambda. \tag{4.16}$$

Let $T(x)$ be defined as the inverse of A_t, with $A_t := \int_0^t \mathfrak{R}(s)ds$, and $A_{T(x)} = x$. Suppose that for some $\delta > 0$ and n large, $T_n(x) \geqslant T(x) + \delta$ or $T_n(x) \leqslant T(x) - \delta$. Then by (4.15), one can show that, for some small $\epsilon > 0$ and n large enough, with high probability

$$x \leqslant \frac{\mathcal{I}_n(T_n(x))}{n} \xrightarrow{p} \lambda\phi^{\mathfrak{R}}(T_n(x)) \geqslant x + \epsilon,$$

or (respectively)

$$x - \frac{1}{n} \geqslant \frac{\mathcal{I}_n(T_n(x))}{n} \xrightarrow{p} \lambda\phi^{\mathfrak{R}}(T_n(x)) \leqslant x - \epsilon.$$

By contradiction, we conclude that with high probability (for large n)

$$T(x) - \delta < T_n(x) < T(x) + \delta.$$

It therefore follows that, by taking δ arbitrarily small, $T_n(x) \xrightarrow{p} T(x)$ as $n \to \infty$. Thus we can conclude that for $0 \leqslant a < b \leqslant \lambda$,

$$T_n(b) - T_n(a) \xrightarrow{p} t_{\mathfrak{R}}(b) - t_{\mathfrak{R}}(a).$$

Now, we proceed to prove (4.5). Note that the limit of the total fraction of revealed ruinous half-edges in the end is equal to $\|\mathfrak{R}\|_{L^1} \wedge \lambda$. Then, using Lemma 4.16, we can transform the reveal process into a death process with an initial number of balls $\mathcal{I}_n^\star$ and i.i.d. lifetimes. Conditioned on $\mathcal{I}_n^\star$, the moments of revealing are simply the order statistics of $\mathcal{I}_n^\star$ i.i.d. random variables with density function

$$f^{(n)}(t) := \frac{\mathcal{R}_n(t)}{\|\mathcal{R}_n\|_{L^1} \wedge (n\lambda^{(n)})} = \frac{\mathfrak{R}(t)}{\|\mathfrak{R}\|_{L^1} \wedge \lambda} + o_p(1), \tag{4.17}$$

for $t \leqslant t_{\mathfrak{R}}(\lambda)$. By using dominated convergence theorem, the above equation implies that

$$\sup_{t\geqslant 0}\Big|\int_0^{t\wedge t_{\mathfrak{R}}(\lambda)} f^{(n)}(s)ds - \frac{\lambda}{\|\mathfrak{R}\|_{L^1} \wedge \lambda}\phi^{\mathfrak{R}}(t)\Big| \xrightarrow{p} 0. \tag{4.18}$$

By the above analysis, it is clear that $Y^{(n)}(t)$ is a pure death process with an initial number of balls $\mathcal{I}_n^\star$ and i.i.d. lifetimes with density $f^{(n)}(t)$, defined by (4.17). Thus, it follows from the Glivenko-Cantelli Theorem that

$$\sup_{t\geqslant 0}\Big|\frac{Y^{(n)}(t)}{\mathcal{I}_n^\star} - \int_0^{t\wedge t_{\mathfrak{R}}(\lambda)} f^{(n)}(s)ds\Big| \xrightarrow{p} 0. \tag{4.19}$$

Combining (4.16), (4.18) and (4.19), we obtain that

$$\begin{aligned}
&\sup_{t\geqslant 0}\frac{1}{\|\mathfrak{R}\|_{L^1} \wedge \lambda}\Big|\frac{Y^{(n)}(t)}{n} - \int_0^{t\wedge t_{\mathfrak{R}}(\lambda)} \mathfrak{R}(s)ds\Big| \\
&\quad\leqslant \sup_{t\geqslant 0}\Big|\frac{Y^{(n)}(t)}{n(\|\mathfrak{R}\|_{L^1} \wedge \lambda)} - \frac{Y^{(n)}(t)}{\mathcal{I}_n^\star}\Big| + \sup_{t\geqslant 0}\Big|\frac{Y^{(n)}(t)}{\mathcal{I}_n^\star} - \int_0^{t\wedge t_{\mathfrak{R}}(\lambda)} f^{(n)}(s)ds\Big| \\
&\qquad + \sup_{t\geqslant 0}\Big|\int_0^{t\wedge t_{\mathfrak{R}}(\lambda)} f^{(n)}(s)ds - \phi^{\mathfrak{R}}(t)\Big| \xrightarrow{p} 0,
\end{aligned} \tag{4.20}$$

which implies that

$$\sup_{t\geqslant 0}\Big|\frac{Y^{(n)}(t)}{n} - \int_0^{t\wedge t_{\mathfrak{R}}(\lambda)} \mathfrak{R}(s)ds\Big|, \tag{4.21}$$

as desired.

4.4.3 Proof of Theorem 4.5

Recall that $\mathcal{I}_n(t)$ represents the number of revealed ruined outgoing half-edges at time t before the stopping time $\tau_n^\star$. According to the construction of the configuration model, all pairs of outgoing and incoming half-edges are chosen uniformly at random. Even if the contagion stops before all half-edges are revealed, we continue the reveal process until the end for the sake of our analysis. This will not change the process before $\tau_n^\star$ and hence will not affect our results.

Based on Lemma 4.16, conditioned on the total number of final revealed ruinous outgoing half-edges, $\mathcal{I}_n^\star$, we can consider the contagion process as follows: There are $\mathcal{I}_n^\star$ ruinous half-edges in total, each incoming half-edge pairs with a ruinous outgoing half-edge with probability $\mathcal{I}_n^\star/n\lambda^{(n)}$ independently (see Remark 4.17 below). If it pairs, this occurs after a random time from the start, with density $f^{(n)}(t)$, defined in (4.17).

Remark 4.17. *Note that the following two events are not generally equivalent:*

(i) Each incoming half-edge is paired with a ruinous outgoing half-edge with probability $\mathcal{I}_n^\star/n\lambda^{(n)}$ independently;

(ii) All pairs of incoming and outgoing half-edges are matched uniformly at random, with a total of $\mathcal{I}_n^\star$ ruinous outgoing half-edges.

However, as n grows, the number of ruinous half-edges in (i) will approach $\mathcal{I}_n^\star$ with high probability, due to the strong law of large numbers. The total number of ruinous half-edges in (i) would be $\mathcal{I}_n^\star + o_p(n)$. This slight deviation does not affect the limit results.

Let us define

$$t_{\mathfrak{R}}^{(n)}(\lambda) := \inf\{t \geqslant 0 : \frac{1}{n}\int_0^t \mathcal{R}_n(s)ds \geqslant \lambda\}.$$

Denote $\mathbb{P}_x^n$ as the probability measure generated by the vectors $\boldsymbol{L}_x$ and $\boldsymbol{U}_x^{\mathfrak{R},(n)}$, where $\boldsymbol{U}_x^{\mathfrak{R},(n)}$ is defined similarly as $\boldsymbol{U}_x^{\mathfrak{R}}$ in Claim 4.12 with the distribution

$$\mathbb{P}(U_{x,i}^{\mathfrak{R},(n)} \leqslant y) = \frac{\int_0^y \mathcal{R}_n(s)ds}{\int_0^{t_{\mathfrak{R}}^{(n)}(\lambda)} \mathcal{R}_n(s)ds}, \quad \text{for} \quad y \leqslant t_{\mathfrak{R}}^{(n)}(\lambda).$$

Note that $\mathbb{P}_x^n$ is a random measure since it depends on $\mathcal{R}_n$. Let $N_x^{(n)}$ denote the number of type x agents in $\mathcal{G}^{(n)}(\mathbf{d}_n^+, \mathbf{d}_n^-)$. By the analysis at the beginning of the proof, we know that, uniformly for all incoming half-edges, the probability of pairing with a ruinous half-edge before time t is

$$\phi_n^{\mathfrak{R}}(t) := \frac{\mathcal{I}_n^\star \int_0^{t\wedge t_{\mathfrak{R}}^{(n)}(\lambda)} \mathcal{R}_n(s)ds}{n\lambda^{(n)} \int_0^{t_{\mathfrak{R}}^{(n)}(\lambda)} \mathcal{R}_n(s)ds},$$

and the probability of never pairing with a ruinous half-edge is $\frac{n\lambda^{(n)} - \mathcal{I}_n^\star}{n\lambda^{(n)}}$.

Under the probability measure $\mathbb{P}^n_x$, the probability of there being θ loss arrivals before time t is given by $b(d^+_x, \phi^{\mathfrak{R}}_n(t), \theta)$. Given θ loss arrivals before time t, the probability of a type x agent being solvent at time t is

$$\mathbb{P}^n_x\big(\tau_{x,0}(\epsilon_x, \boldsymbol{L}_x) = 0, U^{\mathfrak{R},t,(n)}_{(1)} > \tau_{x,1}(\epsilon_x, \boldsymbol{L}_x), U^{\mathfrak{R},t,(n)}_{(2)} > \tau_{x,2}(\epsilon_x, \boldsymbol{L}_x), \ldots, U^{\mathfrak{R},t,(n)}_{(\theta)} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{L}_x)\big),$$

where $U^{\mathfrak{R},t,(n)}_{(1)}, U^{\mathfrak{R},t,(n)}_{(2)}, \ldots, U^{\mathfrak{R},t,(n)}_{(\theta)}$ are the order statistics of θ i.i.d. random variable $\big(U^{\mathfrak{R},t,(n)}_i\big)_{i=1}^{\theta}$ with distribution

$$\mathbb{P}(U^{\mathfrak{R},t,(n)} \leqslant y) = \frac{\int_0^y \mathcal{R}_n(s)ds}{\int_0^t \mathcal{R}_n(s)ds}, \qquad y \leqslant t.$$

We define $\mathcal{S}^{\mathfrak{R},(n)}_{x,\theta}(t)$ as following:

$$\mathcal{S}^{\mathfrak{R},(n)}_{x,\theta}(t) := \mathbb{P}^n_x(\tau_{x,0}(\epsilon_x, \boldsymbol{L}_x) = 0, U^{\mathfrak{R},t,(n)}_{(1)} > \tau_{x,1}(\epsilon_x, \boldsymbol{L}_x), \ldots, U^{\mathfrak{R},t,(n)}_{(\theta)} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{L}_x)).$$

By Claim 4.12, for any $t \geqslant 0$, $x \in \mathcal{X}$ and $0 \leqslant \theta \leqslant d^+_x$, the fraction of solvent agents with exactly θ losses absorbed before time t, namely $S^{(n)}_{x,\theta}(t)/N^{(n)}_x$ is the mapping $\widetilde{\mathbb{P}}^n_x H_{x,\theta,t}$ with respect to the empirical measure of $\mathbb{P}^n_x$, denoted by $\widetilde{\mathbb{P}}^n_x$. By Lemma 4.15, we have for any stopping time $\tau_n \leqslant \tau^\star_n$,

$$\sup_{t \leqslant \tau_n} \Big| \frac{S^{(n)}_{x,\theta}(t)}{N^{(n)}_x} - b(d^+_x, \phi^{\mathfrak{R}}_n(t), \theta)\mathcal{S}^{\mathfrak{R},(n)}_{x,\theta}(t) \Big| \xrightarrow{p} 0. \tag{4.22}$$

Combine the above analysis, clearly we have $\mathbb{P}^n_x H_{x,\theta,t} = b(d^+_x, \phi^{\mathfrak{R}}_n(t), \theta)\mathcal{S}^{\mathfrak{R},(n)}_{x,\theta}(t)$. Notice that, as $n \to \infty$, we have for any $t \geqslant 0$,

$$\phi^{\mathfrak{R}}_n(t) = \frac{\mathcal{I}^\star_n \int_0^{t \wedge t^{(n)}_{\mathfrak{R}}(\lambda)} \mathcal{R}_n(s)ds}{n\lambda^{(n)} \int_0^{t^{(n)}_{\mathfrak{R}}(\lambda)} \mathcal{R}_n(s)ds} \xrightarrow{p} \frac{\int_0^{t \wedge t_{\mathfrak{R}}(\lambda)} \mathfrak{R}(s)ds}{\lambda} = \phi^{\mathfrak{R}}(t),$$

and

$$\frac{n\lambda^{(n)} - \mathcal{I}^\star_n}{n\lambda^{(n)}} \xrightarrow{p} \frac{\lambda - (\|\mathfrak{R}\|_{L^1} \wedge \lambda)}{\lambda}.$$

Combining with dominated convergence theorem, the above two equations give that

$$\sup_{t \geqslant 0} |\mathbb{P}^n_x H_{x,\theta,t} - \mathbb{P}_x H_{x,\theta,t}| \xrightarrow{p} 0. \tag{4.23}$$

Hence by (4.22) and (4.23), we obtain

$$\sup_{t \leqslant \tau_n} \Big| \frac{S^{(n)}_{x,\theta}(t)}{N^{(n)}_x} - b(d^+_x, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t) \Big| \xrightarrow{p} 0.$$

Note that by Assumption 4.1, $N^{(n)}_x/n \longrightarrow \mu_x$ as $n \longrightarrow \infty$, for all $x \in \mathcal{X}$, so we have

$$\sup_{t \leqslant \tau_n} \Big| \frac{N^{(n)}_x}{n} \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t) - \mu_x \mathcal{S}^{\mathfrak{R}}_{x,\theta}(t) \Big| \to 0.$$

Combine the two equations above, we obtain our first assertion

$$\sup_{t\leqslant\tau_n}\Big|\frac{S^{(n)}_{x,\theta}(t)}{n}-\mu_x b(d^+_x,\phi^{\mathfrak{R}}(t),\theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)\Big|\xrightarrow{p}0. \tag{4.24}$$

Let $\mathcal{X}_K$ be the set of all characteristic $x\in\mathcal{X}$ such that $d^+_x+d^-_x\leqslant K$. Since (by Assumption 4.2) $\lambda\in(0,\infty)$, for arbitrary small $\varepsilon>0$, there exists K_ε such that $\sum_{x\in\mathcal{X}\backslash\mathcal{X}_{K_\varepsilon}}\mu_x(d^+_x+d^-_x)<\varepsilon$. Further, by Assumption 4.2 and dominated convergence,

$$\sum_{x\in\mathcal{X}\backslash\mathcal{X}_{K_\varepsilon}}(d^+_x+d^-_x)N^{(n)}_x/n\longrightarrow\sum_{x\in\mathcal{X}\backslash\mathcal{X}_{K_\varepsilon}}(d^+_x+d^-_x)\mu_x<\varepsilon.$$

Hence for n large enough, we have $\sum_{x\in\mathcal{X}\backslash\mathcal{X}_{K_\varepsilon}}(d^+_x+d^-_x)N^{(n)}_x/n<2\varepsilon$. By (4.24), we obtain

$$\begin{aligned}
&\sup_{t\leqslant\tau_n}\sum_{x\in\mathcal{X}}(d^+_x+d^-_x)\sum_{\theta=0}^{d^+_x}\Big|\frac{S^{(n)}_{x,\theta}(t)}{n}-\mu_x b(d^+_x,\phi^{\mathfrak{R}}(t),\theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)\Big|\\
\leqslant&\sup_{t\leqslant\tau_n}\sum_{x\in\mathcal{X}_{K_\varepsilon}}(d^+_x+d^-_x)\sum_{\theta=0}^{d^+_x}\Big|\frac{S^{(n)}_{x,\theta}(t)}{n}-\mu_x b(d^+_x,\phi^{\mathfrak{R}}(t),\theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)\Big|\\
&+\sup_{t\leqslant\tau_n}\sum_{x\in\mathcal{X}\backslash\mathcal{X}_{K_\varepsilon}}(d^+_x+d^-_x)\sum_{\theta=0}^{d^+_x}\Big|\frac{S^{(n)}_{x,\theta}(t)}{n}-\mu_x b(d^+_x,\phi^{\mathfrak{R}}(t),\theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)\Big|\\
\leqslant&o_p(1)+\sum_{x\in\mathcal{X}\backslash\mathcal{X}_{K_\varepsilon}}(d^+_x+d^-_x)(N^{(n)}_x/n+\mu_x)\leqslant o_p(1)+3\varepsilon.
\end{aligned}$$

By taking ε arbitrarily small, it follows that

$$\sup_{t\leqslant\tau_n}\sum_{x\in\mathcal{X}}(d^+_x+d^-_x)\Big|\frac{S^{(n)}_{x,\theta}(t)}{n}-\mu_x b(d^+_x,\phi^{\mathfrak{R}}(t),\theta)\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t)\Big|\xrightarrow{p}0. \tag{4.25}$$

Note that the total number of solvent agents at time t satisfies

$$S^{(n)}(t)=\sum_{x\in\mathcal{X}}\sum_{\theta=0}^{d^+_x}S^{(n)}_{x,\theta}(t),$$

which is dominated by $\sum_{x\in\mathcal{X}}(d^+_x+d^-_x)\sum_{\theta=0}^{d^+_x}S^{(n)}_{x,\theta}(t)$. Then, by the convergence results (4.24) and (4.25), we obtain

$$\sup_{t\leqslant\tau_n}\Big|\frac{S^{(n)}(t)}{n}-f_S(t)\Big|\xrightarrow{p}0.$$

Further, from $D^{(n)}(t)=n-S^{(n)}(t)$, the number of ruined agents at time t also satisfies

$$\sup_{t\leqslant\tau_n}\Big|\frac{D^{(n)}(t)}{n}-f_D(t)\Big|\xrightarrow{p}0.$$

Finally, the total number of remaining ruinous outgoing half-edges at time t is given by

$$W_n(t) = n\lambda^{(n)} - \mathcal{I}_n(t) - \sum_{x\in\mathcal{X}} \sum_{\theta=0}^{d_x^+} d_x^- S_{x,\theta}^{(n)}(t).$$

The same argument and (4.20) imply that

$$\sup_{t\leqslant\tau_n} \Big|\frac{W_n(t)}{n} - f_W(e^{-t})\Big| \xrightarrow{p} 0.$$

This completes the proof of Theorem 4.5.

4.4.4 Proof of Theorem 4.7

Similar to Example 4.3, for any $x \in \mathcal{X}$ and a realization of loss sequence $\boldsymbol{\ell}_x$ and shock ϵ_x, we get

$$P_{x,\theta}^{\mathfrak{R}}(t,\epsilon_x,\boldsymbol{\ell}_x) := \mathbb{P}(\tau_{x,0}(\epsilon_x,\boldsymbol{\ell}_x) = 0, U_{(1)}^{\mathfrak{R},t} > \tau_{x,1}(\epsilon_x,\boldsymbol{\ell}_x),\ldots,U_{(\theta)}^{\mathfrak{R},t} > \tau_{x,\theta}(\epsilon_x,\boldsymbol{\ell}_x)) = \mathbb{1}\{\theta < \delta_x(\epsilon_x,\boldsymbol{\ell}_x)\},$$

where $\delta_x(\epsilon_x,\boldsymbol{\ell}_x) := \inf\{\theta = 0,\ldots,d_x^+ : \gamma_x(1-\epsilon_x) - \delta_x < \ell_{x,1} + \cdots + \ell_{x,\theta}\}$ (by convention we set $\delta_x(\epsilon_x,\boldsymbol{\ell}_x) := \infty$ if there is not such a threshold θ). Hence, from the definition of the default threshold distribution $q_{x,\theta}$, $\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t)$ simplifies to

$$\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t) = \sum_{\delta=\theta+1}^{d_x^+} q_{x,\delta} + \bar{q}_x.$$

By applying Theorem 4.5, the first claim is established. Observing that

$$b(d_x^+, \phi^{\mathfrak{R}}(t), \theta) = b(d_x^+, 1-\phi^{\mathfrak{R}}(t), d_x^+ - \theta),$$

we can rearrange the order of the sums to obtain

$$\begin{aligned}\sum_{\theta=1}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t) &= \sum_{\theta=1}^{d_x^+}\Big[\sum_{\delta=\theta+1}^{d_x^+} q_{x,\delta} + \bar{q}_x\Big] b(d_x^+, 1-\phi^{\mathfrak{R}}(t), d_x^+ - \theta)\\ &= \bar{q}_x + \sum_{\theta=1}^{d_x^+} q_{x,\theta}\beta\big(d_x^+, 1-\phi^{\mathfrak{R}}(t), d_x^+ - \theta + 1\big).\end{aligned}$$

This leads to the conclusion for the limit function for $\widehat{f}_S^{\mathfrak{R}}(t)$. Through similar arguments and calculation, the limit functions for the other cases can also be derived.

4.4.5 Proof of Lemma 4.8

Recall that

$$f_W^{\mathfrak{R}}(t) := \lambda(1-\phi^{\mathfrak{R}}(t)) - \sum_{x\in\mathcal{X}} \mu_x d_x^- \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t).$$

Consider a constant $t_1 \in (0, t_{\mathfrak{R}}^\star)$. Due to the continuity of $f_W^{\mathfrak{R}}(t)$ on the interval $[0,\infty)$, it follows that $f_W^{\mathfrak{R}}(t) > 0$ on $[0, t_1)$. Thus, there exists a constant $C_1 > 0$ such that $f_W^{\mathfrak{R}}(t) \geqslant C_1$ for all $t \leqslant t_1$.

Since $W_n(\tau_n^\star) = 0$, if $\tau_n^\star \leqslant t_1$, then we have $W_n(\tau_n^\star)/n - f_W^{\mathfrak{R}}(\tau_n^\star) \leqslant -C_1$. But on the other hand, according to Theorem 4.5, we have

$$\sup_{t\leqslant\tau_n^\star} |\frac{W_n(t)}{n} - f_W^{\mathfrak{R}}(t)| \xrightarrow{p} 0.$$

This is a contradiction. Hence, it must be the case that $\mathbb{P}(\tau_n^\star \leqslant t_1) \longrightarrow 0$, as $n \longrightarrow \infty$. In the case where $t_{\mathfrak{R}}^\star = \infty$, we can choose any finite t_1 arbitrarily large, which implies that $\tau_n^\star \xrightarrow{p} \infty$.

We now consider the other scenario. Fix a constant $t_2 \in (t_{\mathfrak{R}}^\star, t_{\mathfrak{R}}^\star + \varepsilon)$. Using a similar argument as above, we can show that there exists some constant $C_2 > 0$ such that $W_n(\tau_n^\star)/n - f_W^{\mathfrak{R}}(\tau_n^\star) \geqslant C_2$ if $\tau_n^\star \geqslant t_2$. Therefore, $\mathbb{P}(\tau_n^\star \geqslant t_2) \longrightarrow 0$ as $n \longrightarrow \infty$. As t_1 and t_2 are arbitrary, letting both t_1 and t_2 tend to $t_{\mathfrak{R}}^\star$, we have $\tau_n^\star \xrightarrow{p} t_{\mathfrak{R}}^\star$.

4.4.6 Proof of Theorem 4.9

By using Lemma 4.8, we have that the stopping time $\tau_n^\star$ converges to $t_{\mathfrak{R}}^\star$ in probability. In combination with the continuity of $\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t)$ in t and Theorem 4.5,

$$\sup_{t\leqslant\tau_n^\star} \Big|\frac{D_x^{(n)}(t)}{n} - \mu_x\big(1 - \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t)\big)\Big| \xrightarrow{p} 0,$$

as $n \to \infty$. Besides, If $\int_0^{t_{\mathfrak{R}}^\star} |\mathfrak{R}(s)| ds = \lambda$, that means at the stopping time $\tau_n^\star$, we have revealed almost all the outgoing half-edges. Thus the number of defaults must be $n - o_p(n)$. Moreover, by Lemma 4.8, we have $\tau_n^\star \xrightarrow{p} \infty$. Notice that $\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t)$ is non-increasing and can not be smaller than zero. Thus there exists a limit for $\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t)$ when $t \to \infty$. Note also that, if $\|\mathfrak{R}\|_{L^1} < \lambda$, $\phi^{\mathfrak{R}}(t) \to \|\mathfrak{R}\|_{L^1}/\lambda$ as $t \to \infty$. Then the theorem follows by Lemma 4.8 and Theorem 4.5.

4.4.7 Proof of Theorem 4.10

Proof of (i). The proof of (i) will be divided into two parts.

Lipschitz continuity. Let $\|\cdot\|_t$ denote the truncated $\mathbb{L}^1$ norm up to t, i.e., for a function f defined on $\mathbb{R}^+$, $\|f\|_t = \int_0^t |f(s)| ds$. The first step is to prove that Ψ has some special Lipschitz continuity with respect to the first parameter. To be precise, for two different $\mathfrak{R}_1, \mathfrak{R}_2 \in \mathbb{L}_\lambda(\mathbb{R}^+)$ and all $t > 0$, there exists a constant C_0, such that

$$|\Psi(\mathfrak{R}_1)(t) - \Psi(\mathfrak{R}_2)(t)| \leqslant C_0 \|\mathfrak{R}_1 - \mathfrak{R}_2\|_t. \tag{4.26}$$

For the first term, it is clear that

$$\lambda|\phi^{\mathfrak{R}_1}(t) - \phi^{\mathfrak{R}_2}(t)| \leqslant \int_0^t |\mathfrak{R}_1(s) - \mathfrak{R}_2(s)| ds.$$

We now analyze the difference $|\mathcal{S}^{\mathfrak{R}_1}_{x,\theta}(t) - \mathcal{S}^{\mathfrak{R}_2}_{x,\theta}(t)|$. Let us recall the definitions of

$$P^{\mathfrak{R}}_{x,\theta}(t, \epsilon_x, \boldsymbol{\ell}_x) := \mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{\ell}_x) = 0, U^{\mathfrak{R},t}_{(1)} > \tau_{x,1}(\epsilon_x, \boldsymbol{\ell}_x), \dots, U^{\mathfrak{R},t}_{(\theta)} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x)),$$

and

$$\mathcal{S}^{\mathfrak{R}}_{x,\theta}(t) = \mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{L}_x) = 0, U^{\mathfrak{R},t}_{(1)} > \tau_{x,1}(\epsilon_x, \boldsymbol{L}_x), \dots, U^{\mathfrak{R},t}_{(\theta)} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{L}_x)).$$

For each realization of loss sequence $\boldsymbol{\ell}_x$ and shock ϵ_x, let $\mathcal{A}_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x)$ be the set of all $(u_1, u_2, \dots, u_\theta)$ such that the rearranged sequence $(u_{(1)}, u_{(2)}, \dots, u_{(\theta)})$ in increasing order satisfies

$$u_{(1)} > \tau_{x,1}(\epsilon_x, \boldsymbol{\ell}_x), \dots, u_{(\theta)} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x).$$

Let $\mathfrak{R}^t(s) = \mathfrak{R}(s)/\|\mathfrak{R}\|_t$. We have

$$\begin{aligned}
&\int_{\mathcal{A}_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x)} \mathfrak{R}^t_1(s_1) \cdots \mathfrak{R}^t_1(s_\theta) ds_1 \cdots ds_\theta - \int_{\mathcal{A}_{x,\theta}(\boldsymbol{\ell}_x)} \mathfrak{R}^t_2(s_1) \cdots \mathfrak{R}^t_2(s_\theta) ds_1 \cdots ds_\theta \\
\leqslant &\int_{\mathcal{A}_{x,\theta}(\boldsymbol{\ell}_x)} |\mathfrak{R}^t_1(s_1) \cdots \mathfrak{R}^t_1(s_\theta) - \mathfrak{R}^t_2(s_1) \cdots \mathfrak{R}^t_2(s_\theta)| ds_1 \cdots ds_\theta \\
\leqslant &\int_{[0,t]^\theta} |\mathfrak{R}^t_1(s_1) \cdots \mathfrak{R}^t_1(s_\theta) - \mathfrak{R}^t_2(s_1) \cdots \mathfrak{R}^t_2(s_\theta)| ds_1 \cdots ds_\theta.
\end{aligned}$$

By adding and subtracting terms, it is easy to show that

$$\int_{[0,t]^\theta} |\mathfrak{R}^t_1(s_1) \cdots \mathfrak{R}^t_1(s_\theta) - \mathfrak{R}^t_2(s_1) \cdots \mathfrak{R}^t_2(s_\theta)| ds_1 \cdots ds_\theta \leqslant \theta \|\mathfrak{R}^t_1 - \mathfrak{R}^t_2\|_t.$$

Notice that there exists a constant C_1 large enough such that

$$\begin{aligned}
\|\mathfrak{R}^t_1 - \mathfrak{R}^t_2\|_t \leqslant & \left\|\frac{\mathfrak{R}_1}{\|\mathfrak{R}_1\|_t} - \frac{\mathfrak{R}_2}{\|\mathfrak{R}_1\|_t}\right\|_t + \left\|\frac{\mathfrak{R}_2}{\|\mathfrak{R}_1\|_t} - \frac{\mathfrak{R}_2}{\|\mathfrak{R}_2\|_t}\right\|_t \\
\leqslant & \frac{\|\mathfrak{R}_1 - \mathfrak{R}_2\|_t}{\|\mathfrak{R}_1\|_t} + \frac{\|\mathfrak{R}_2\|_t - \|\mathfrak{R}_1\|_t}{\|\mathfrak{R}_1\|_t \|\mathfrak{R}_2\|_t} \|\mathfrak{R}_2\|_t \\
\leqslant & C_1 \|\mathfrak{R}_1 - \mathfrak{R}_2\|_t.
\end{aligned}$$

Therefore, for any realization of loss sequence $\boldsymbol{\ell}_x$ and shock ϵ_x, we have

$$\mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{\ell}_x) = 0, U_{(1)}^{\mathfrak{R},t} > \tau_{x,1}(\epsilon_x, \boldsymbol{\ell}_x), \dots, U_{(\theta)}^{\mathfrak{R},t} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{\ell}_x)) \leqslant C_1\theta\|\mathfrak{R}_1 - \mathfrak{R}_2\|_t.$$

It therefore follows that

$$\mathbb{P}(\tau_{x,0}(\epsilon_x, \boldsymbol{L}_x) = 0, U_{(1)}^{\mathfrak{R},t} > \tau_{x,1}(\epsilon_x, \boldsymbol{L}_x), \dots, U_{(\theta)}^{\mathfrak{R},t} > \tau_{x,\theta}(\epsilon_x, \boldsymbol{L}_x)) \leqslant C_1\theta\|\mathfrak{R}_1 - \mathfrak{R}_2\|_t. \tag{4.27}$$

On the other hand, by elementary calculation for the derivative of $b(d_x^+, y, \theta)$ with respect to y,

$$\frac{\partial b(d_x^+, y, \theta)}{\partial y} = d_x^+ b(d_x^+ - 1, y, \theta - 1) - d_x^+ b(d_x^+ - 1, y, \theta),$$

and

$$\frac{\partial b(d_x^+, y, 0)}{\partial y} = -d_x^+ b(d_x^+ - 1, y, 0), \qquad \frac{\partial b(d_x^+, y, d_x^+)}{\partial y} = d_x^+ b(d_x^+ - 1, y, d_x^+ - 1).$$

For each $y \in [0,1]$, there exists a threshold $\bar{\theta}_y$ such that $\partial b(d_x^+, y, \theta)/\partial y \leqslant 0$ for $\theta \leqslant \bar{\theta}_y$ and $\partial b(d_x^+, y, \theta)/\partial y \geqslant 0$ for $\theta > \bar{\theta}_y$. Therefore we have for all $x \in \mathcal{X}$,

$$\frac{\partial \sum_{\theta=0}^{d_x^+} a_{x,\theta} b(d_x^+, y, \theta)}{\partial y} \leqslant 2d_x^+ b(d_x^+ - 1, y, \bar{\theta}_y) \leqslant 2d_x^+, \tag{4.28}$$

if $a_{x,\theta}$ is a constant coefficient satisfying $0 \leqslant a_{x,\theta} \leqslant 1$ for all (x, θ).

Then by adding and subtracting terms and combining (4.27) and (4.28), we conclude that there exists a constant C_2 such that for each $x \in \mathcal{X}$,

$$\begin{aligned}
&\Big|\sum_{\theta=0}^{d_x^+}(b(d_x^+, \phi^{\mathfrak{R}_1}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}_1}(t) - b(d_x^+, \phi^{\mathfrak{R}_2}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}_2}(t))\Big| \\
&\qquad \leqslant \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}_2}(t), \theta)(C_1\theta)\|\mathfrak{R}_1 - \mathfrak{R}_2\|_t + 2d_x^+|\phi^{\mathfrak{R}_1}(t) - \phi^{\mathfrak{R}_2}(t)| \\
&\qquad \leqslant \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}_2}(t), \theta)C_1 d_x^+\|\mathfrak{R}_1 - \mathfrak{R}_2\|_t + 2\frac{d_x^+}{\lambda}\|\mathfrak{R}_1 - \mathfrak{R}_2\|_t \leqslant C_2 d_x^+\|\mathfrak{R}_1 - \mathfrak{R}_2\|_t.
\end{aligned}$$

It therefore follows from Assumption 4.4 that there exists a constant C_0 such that

$$\sum_{x\in\mathcal{X}} \mu_x q_x d_x^- \Big|\sum_{\theta=0}^{d_x^+}(\mathcal{S}_{x,\theta}^{\mathfrak{R}_1}(t) - \mathcal{S}_{x,\theta}^{\mathfrak{R}_2}(t))\Big| \leqslant C_1 \sum_{x\in\mathcal{X}} \mu_x q_x d_x^- d_x^+\|\mathfrak{R}_1 - \mathfrak{R}_2\|_t \leqslant C_0\|\mathfrak{R}_1 - \mathfrak{R}_2\|_t,$$

and hence (4.26) holds.

Existence and uniqueness. We show existence using a standard iterative procedure. To prove the existence and uniqueness on the non-negative real numbers $\mathbb{R}^+$, it is sufficient to prove them on the interval $[0,T]$ for any arbitrary $T > 0$. Specifically, let $h^{(0)}(t) = \beta \sum_{x \in \mathcal{X}} \mu_x d_x^-(1 - q_{x,0}) - t$, and define

$$h^{(n)}(t) = \beta \Psi(h^{(n-1)})(t). \tag{4.29}$$

By (4.26), we have (for $v \in \mathbb{R}_+$)

$$|h^{(n)}(v) - h^{(n-1)}(v)| \leqslant \beta C_0 \|h^{(n-1)} - h^{(n-2)}\|_v.$$

Integrating the above formula from $v = 0$ to $v = t$ gives us

$$\|h^{(n)} - h^{(n-1)}\|_t \leqslant \beta C_0 \int_0^t \|h^{(n-1)} - h^{(n-2)}\|_v dv. \tag{4.30}$$

Clearly $\|h^{(1)} - h^{(0)}\|_T \leqslant C$ for some constant C. Then by iterating the above formula (4.30), we have

$$\|h^{(n)} - h^{(n-1)}\|_t \leqslant \frac{C}{(n-1)!} (\beta C_0)^{n-1} t^{n-1}.$$

Moreover, for some constant $C(\beta, K, T)$ depending on β, K and T, the infinite sum satisfies

$$\sum_{n=1}^{\infty} \|h^{(n)} - h^{(n-1)}\|_T \leqslant C(\beta, C_0, T) e^{C(\beta, C_0, T)}.$$

Therefore the series $h^{(n)}(t)$ converges in the $L^1([0,T])$ space to a limit $\mathfrak{F}(t)$. Due to the continuity of the limit, we can conclude the existence of a solution.

To prove uniqueness, suppose there exists two different solutions $h_1(t)$ and $h_2(t)$ satisfying the fixed point equation. We have

$$\|h_1 - h_2\|_T \leqslant C_0 \int_0^T \|h_1 - h_2\|_v dv.$$

Since the function $\|h_1 - h_2\|_t$ is bounded on $[0,T]$ and positive, the Gronwall Lemma implies that $\|h_1 - h_2\|_t = 0$ on $[0,T]$. Since T is arbitrary and the solution is continuous, uniqueness follows.

Proof of (ii)**.** By the construction of our networked risk processes, the loss reveal function $\mathcal{R}_n(t)$ is equal to the $\beta W_n(t)$. Suppose that the limit process of $W_n(t)$ exists and satisfies the conditions outlined in Theorem 4.5. From the definition of $f_W^{\mathfrak{R}^\star}(t)$, we have

$$\frac{\mathfrak{R}^\star(t)}{\beta} = f_W^{\mathfrak{R}^\star}(t) = \lambda \phi^{\mathfrak{R}^\star}(t) - \sum_{x \in \mathcal{X}} \mu_x d_x^- \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}^\star}(t), \theta) \mathcal{S}_{x,\theta}^{\mathfrak{R}^\star}(t).$$

The existence and uniqueness of the solution to this fixed point equation were proved in point (i). As a result, by Theorem 4.5, the unique solution $\mathfrak{R}^\star$ of the fixed point equation in (i) is the limit process of $\beta W_n(t)$ and we have

$$\sup_{t \leqslant \tau_n^\star} \left| \frac{W_n(t)}{n} - \frac{\mathfrak{R}^\star(t)}{\beta} \right| \xrightarrow{p} 0.$$

The remaining limit results follow directly from Theorem 4.5.

The proof of Theorem 4.9 is now complete.

4.4.8 Proof of Lemma 4.11

Let $\alpha_n n$ denote the number of initially ruined outgoing half-edges. Note that, since $\mathfrak{R}^\star(0) = \beta \sum_{x\in\mathcal{X}} \mu_x d_x^-(1-q_{x,0}) > 0$, we have $\liminf_n \alpha_n > 0$. Let Λ_n denote the time required to reveal all the $\alpha_n n$ initially ruined outgoing half-edges without incurring any new ruined agents, i.e.

$$\Lambda_n = T_1^{(n)} + T_2^{(n)} + \cdots + T_{\alpha_n n}^{(n)},$$

where $T_k^{(n)}$ is the time duration of the k-th reveal and is an exponential random variable with parameter $\alpha_n n - k + 1$, and they are independent. If no new ruins incurred at each step, $\tau_n^\star$ will attain the smallest possible value stochastically, namely $\tau_n^\star \geqslant_{\text{st}} \Lambda_n$. It is therefore sufficient to prove that $\Lambda_n > (1-\epsilon)\log n$ with high probability for any $\epsilon > 0$. By Markov's inequality we have

$$\mathbb{P}(\Lambda_n \leqslant (1-\epsilon)\log n) = \mathbb{P}(e^{-\Lambda_n} \geqslant e^{-(1-\epsilon)\log n}) \leqslant e^{(1-\epsilon)\log n} \prod_{k=1}^{\alpha_n n} \mathbb{E} e^{-T_k^{(n)}}.$$

Since $\mathbb{E}e^{-T_k^{(n)}} = (\alpha_n n - k + 1)/(\alpha_n n - k)$, we have

$$\mathbb{P}(\Lambda_n \leqslant (1-\epsilon)\log n) \leqslant \exp\{(1-\epsilon)\log n - \sum_{k=1}^{\alpha_n n} \log(1 + \frac{1}{k})\}.$$

By applying Taylor expansion to $\log(1+1/k)$, it follows that

$$\begin{aligned} \mathbb{P}(\Lambda_n \leqslant (1-\epsilon)\log n) &\leqslant \exp\{(1-\epsilon)\log n - \log(\alpha_n n) - o(\log n)\} \\ &\leqslant \exp\{(1-\epsilon)\log n - \log n - \log \alpha_n - o(\log n) = O(n^{-\epsilon}). \end{aligned}$$

We thus have $\Lambda_n > (1-\epsilon)\log n$ with high probability for any $\epsilon > 0$. The proof is complete.

4.5 Complex Networked Risk Processes

So far we have assumed that the external debt is a constant function for each agent. It will be interesting to extend the model by considering a dynamics for this external debt that is itself like in the classical Cramér-Lundberg model. Namely, we assume that the external debt for agent $i \in [n]$ follows $\delta_i(t) = \sum_{j=1}^{N_i(t)} \zeta_i^{(j)}$, where $N_i(t)$ is a Poisson process with intensity β_i and the claim sizes $\{\zeta_i^{(j)}\}_{j=1}^\infty$ are i.i.d. distributed random variables with distribution G_i with finite positive mean and variance.

The risk process for the capital of agent $i \in [n]$ with network interactions $\mathcal{G}_n$ follows

$$C_i(t) := \gamma_i(1-\epsilon_i) + \alpha_i(t) - \sum_{j=1}^{N_i(t)} \zeta_i^{(j)} - \sum_{j\in[n]:j\to i} L_{ji}\mathbb{1}\{\tau_j + T_{ji} \leqslant t\}, \tag{4.31}$$

where, similar to (4.3) we assume that $T_{ji} \sim \text{Exp}(\beta)$ are i.i.d. exponentially distributed with some parameter $\beta > 0$ for all $i, j \in [n]$.

Consider the node classification of Section 4.2.2 and assume that for all agents of the same type $x \in \mathcal{X}$, the associated external risk process has the same features. Therefore, agents of type x have the same claim distribution, denoted by G_x and external claim arrival intensity denoted by β_x.

Let us define the external risk process for an agent i of type $x \in \mathcal{X}$ with initial capital u by

$$C_x^{\mathrm{EX}}(u,t) := u + \alpha_x(t) - \sum_{j=1}^{N_x(t)} \zeta_i^{(j)},$$

where $N_x(t)$ is a Poisson process with intensity β_x and the claim sizes $\{\zeta_i^{(j)}\}_{j=1}^{\infty}$ are i.i.d. distributed random variables with distribution G_x and mean $\bar{\zeta}_x > 0$. Similarly, for the network $\mathcal{G}_n$ and initial type-dependent capital vector $\mathbf{u} = (u_x, x \in \mathcal{X})$, we define the internal risk process for agent i by

$$C_x^{\mathrm{IN}}(\mathbf{u},t) := u_x + \alpha_x(t) - \sum_{j \in [n]: j \to i} L_{ji} \mathbb{1}\{\tau_j + T_{ji} \leqslant t\},$$

where losses L_{ij} are i.i.d. random variables with distribution F_x and $T_{ji} \sim \mathrm{Exp}(\beta)$ are i.i.d. exponentially distributed.

For simplicity we assume that ϵ_x is a constant and $\alpha_x(t) = \alpha_x t$ in this section.

Let us denote by

$$\psi_x^{\mathrm{EX}}(u,t) = \mathbb{P}(C_x^{\mathrm{EX}}(u,s) \leqslant 0, \text{ for some } s \leqslant t)$$

and

$$\psi_x^{\mathrm{IN}}(\mathbf{u},t) = \mathbb{P}(C_x^{\mathrm{IN}}(\mathbf{u},s) \leqslant 0, \text{ for some } s \leqslant t),$$

the ruin probabilities for the external and respectively internal risk processes of an agent of type x. The ruin probability for the external process is a well studied problem whose solution we review below. The ruin probability for the internal process is given by Theorem 4.10 in the limit when n is large and converges to

$$\psi_x^{\mathrm{IN}}(\mathbf{u},t) = 1 - \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}^\star}(t), \theta) \mathcal{S}_{x,\theta}^{\mathfrak{R}^\star}(\infty) + o_p(1).$$

Note that the parameters (and in particular the fraction of initial ruined agents) depend on the initial capital levels $\mathbf{u}$, which reflect any initial shock and external debt of our baseline process in Section 4.3.2. We will use these ruin probabilities to provide an upper bound and lower bound on the ruin probability for agent i of type x given by

$$\psi_x(t) := \mathbb{P}(C_x(s) \leqslant 0, \text{ for some } s \leqslant t),$$

where

$$C_x(t) := \gamma_x(1-\epsilon_x) + \alpha_x t - \sum_{j=1}^{N_x(t)} \zeta_i^{(j)} - \sum_{j \in [n]: j \to i} L_{ji} \mathbb{1}\{\tau_j + T_{ji} \leqslant t\}.$$

It is known (see e.g. [35, 115]) that whenever $\beta_x\bar{\zeta}_x > \alpha_x$, we have $\psi_x^{\mathrm{EX}}(u,\infty) = 1$ for all $u \in \mathbb{R}$ and whenever $\beta_x\bar{\zeta}_x < \alpha_x$, the final ruin probability can be computed using the famous Pollaczek–Khinchine formula as

$$\psi_x^{\mathrm{EX}}(u,\infty) = \left(1 - \frac{\beta_x\bar{\zeta}_x}{\alpha_x}\right)\sum_{k=0}^{\infty}\left(\frac{\beta_x\bar{\zeta}_x}{\alpha_x}\right)^k\left(1 - \widehat{G}_x^{*k}(\gamma)\right), \tag{4.32}$$

where

$$\widehat{G}_x(u) = \frac{1}{\bar{\zeta}_x}\int_0^u \big(1 - G_x(u)\big)du,$$

and the operator $(\cdot)^{*k}$ denotes the k-fold convolution.

Furthermore, for the finite time horizon ruin probability $\psi_x^{\mathrm{EX}}(u,t)$, we have the Seal-type formula, see e.g. [164, Proposition 3.4], for any $t > 0$ and any $x \in \mathcal{X}$,

$$\begin{aligned}\psi_x^{\mathrm{EX}}(\beta_x u,t) =& 1 - e^{-\beta_x t} - \int_0^{u+t} f_t(z)dz + \int_u^{u+t} e^{-\beta_x(t+u-z)}f_{z-u}(z)dz \\ &+ \int_u^{u+t} f_{z-u}(z)\Big(\int_z^{u+t}\frac{t+u-y}{t+u-z}f_{t+u-z}(y-z)dy\Big)dz,\end{aligned}$$

where $f_t^x(\cdot)$ denotes the density of $\sum_{j=1}^{N_x(t)}\zeta_i^{(j)}/\beta_x$ on $(0,\infty)$.

The following theorem provides upper and lower bounds for the complex ruin probability $\psi_x(t)$ by utilizing ψ_x^{IN} and ψ_x^{EX}. We define $\widetilde{\tau}_{x,\theta}$ similarly to $\tau_{x,\theta}$ in Equation (4.7), but with α_x and γ_x replaced by $\alpha_x/2$ and $\gamma_x/2$, respectively. Furthermore, for any $\mathfrak{R} \in \mathbb{L}_\lambda(\mathbb{R}^+)$, we define $\widetilde{\mathcal{S}}_{x,\theta}^{\mathfrak{R}}(t)$ in the same manner as $\mathcal{S}_{x,\theta}^{\mathfrak{R}}(t)$, but with $\tau_{x,\theta}$ replaced by $\widetilde{\tau}_{x,\theta}$.

Theorem 4.18. *Let $\mathfrak{R}^\star$ be the unique solution of the fixed point equation in Theorem 4.10 with an initial value $\mathfrak{R}^\star(0) = \beta\sum_{x\in\mathcal{X}}\mu_x d_x^-(1-q_{x,0})$, and let $\widetilde{\mathfrak{R}}^\star$ be the unique solution of*

$$\widetilde{\mathfrak{R}}^\star(t) = \lambda(1 - \phi^{\widetilde{\mathfrak{R}}^\star}(t)) - \sum_{x\in\mathcal{X}}\mu_x d_x^-\sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\widetilde{\mathfrak{R}}^\star}(t), \theta)\widetilde{\mathcal{S}}_{x,\theta}^{\widetilde{\mathfrak{R}}^\star}(t),$$

with an initial value $\widetilde{\mathfrak{R}}^\star(0) = \sum_{x\in\mathcal{X}}\mu_x d_x^-(1-q_{x,0})$. Also, let $\psi_x^{\mathrm{EX}}(t) := \psi_x^{\mathrm{EX}}(\gamma_x(1-\epsilon_x),t)$ and let $\widetilde{\psi}_x^{\mathrm{EX}}(t)$ be the ruin probability for the external risk process at time t staring with half the initial value $\gamma_x(1-\epsilon_x)/2$ and half the capital growth rate $\alpha_x/2$. Under Assumption 4.1 and Assumption 4.4, for all $x \in \mathcal{X}$ and for any $t > 0$, when n is large enough, we have

$$\psi_x(t) \geqslant 1 - (1 - \psi_x^{\mathrm{EX}}(t))\sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}^\star}(t), \theta)\mathcal{S}_{x,\theta}^{\mathfrak{R}^\star}(t),$$

and

$$\psi_x(t) \leqslant 1 - (1 - \widetilde{\psi}_x^{\mathrm{EX}}(t))\sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\widetilde{\mathfrak{R}}^\star}(t), \theta)\widetilde{\mathcal{S}}_{x,\theta}^{\widetilde{\mathfrak{R}}^\star}(t).$$

Proof. We write all the quantities in their limit forms, since this theorem states the bounds when n is large enough. We define the events

$$A_t := \{C_x^{\mathrm{EX}}(s) > 0, \text{for all } s \leqslant t\},$$

$$B_t := \{C_x^{\mathrm{IN}}(s) > 0, \text{for all } s \leqslant t \text{ with an initial fraction of ruinous half-edges } \sum_{x \in \mathcal{X}} \mu_x d_x^-(1 - q_{x,0})\}$$

and,

$$D_t := \{C_x(s) > 0, \text{for all } s \leqslant t \text{ with an initial fraction of ruinous half-edges } \sum_{x \in \mathcal{X}} \mu_x d_x^-(1 - q_{x,0})\}.$$

Then by the independence of external and internal risk processes, we have

$$\mathbb{P}(D_t) \leqslant \mathbb{P}(A_t \cap B_t) = \mathbb{P}(A_t) \cdot \mathbb{P}(B_t) = (1 - \psi_x^{\mathrm{EX}}(t)) \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}^\star}(t), \theta) \mathcal{S}_{x,\theta}^{\mathfrak{R}^\star}(t).$$

Since $\psi_x(t) = \mathbb{P}(D_t^c)$, then we obtain the lower bound in the theorem. To prove the upper bound, we define the following events

$$\widetilde{A}_t := \{ \sum_{j=1}^{N_i(t)} \zeta_i^{(j)} < \frac{\gamma_x}{2} + \frac{\alpha_x(s)}{2}, \text{for all } s \leqslant t\},$$

$$\widetilde{B}_{T,t} := \{L_{ji} \mathbb{1}\{\tau_j + T_{ji} \leqslant t\} < \frac{\gamma_x}{2} + \frac{\alpha_x(s)}{2}, \text{for all } s \leqslant t \text{ with the loss reveal intenstiy } \widetilde{\mathfrak{R}}^\star\}.$$

Then we have

$$\mathbb{P}(D_t) \geqslant \mathbb{P}(\widetilde{A}_t \cap \widetilde{B}_{T,t}) = \mathbb{P}(\widetilde{A}_t) \cdot \mathbb{P}(\widetilde{B}_{T,t}) = (1 - \widetilde{\psi}_x^{\mathrm{EX}}(t)) \sum_{\theta=0}^{d_x^+} b(d_x^+, \phi^{\mathfrak{R}^\star}(t), \theta) \widetilde{\mathcal{S}}_{x,\theta}^{\widetilde{\mathfrak{R}}^\star_T}(t).$$

Since $\psi_x(t) = \mathbb{P}(D_t^c)$, the upper bound in theorem follows.

□

The general setup of Section 4.3.1 for the internal risk process can be considered where the limit function for the loss reveal intensity function is a given function $\mathfrak{R}$ that satisfies Assumption 4.3. In this case, both the upper and lower bounds in the theorem above still hold, with $\mathfrak{R}^\star$ and $\widetilde{\mathfrak{R}}^\star$ being replaced by $\mathfrak{R}$. The question of finding exact (asymptotic) ruin probabilities in this complex networked risk processes is left for future research.

4.6 Concluding Remarks

We have extended and solved an open problem posed in our prior work [29], namely a multi-dimensional extension to the Cramér-Lundberg risk processes with network-driven losses. We called this the (internal) networked risk processes. In this general case, the losses from the ruined neighbors can have any given distribution. There is a general arrival process which drives the losses stemming from the ruined agents. This model opens the way to integrating two streams of literature, one being on risk models that have outside exogenous losses and one where losses are internal and interconnected.

This is one step forward to study dynamic financial network models and at the same time bringing tractability to multi-dimensional Cramér-Lundberg risk processes. We leave to the future the study of further generalizations. Here we mention only a few of them. In Section 4.5 we introduced a model that allows for exogenous individual external risk processes in addition to the internal networked one. We have provided lower and upper bounds for the ruin probabilities. Finding the exact ruin probabilities in this case is left for future research. Instead of a sparse network structure that underlies the internal loss propagation model, we could consider risk processes in a dense network. In particular, since our results are asymptotic in nature, it seems promising to consider graphons. These have been developed by Lovász et al., see e.g. [67, 68, 170], as a natural continuum limit object for large dense graphs. It would also be interesting to optimize the underlying graphon structure given the profile of external losses. When these external losses follow a general risk process, we can expect to have various types of adaptive graphons that minimize the loss contagion.

The model we consider is flexible enough to incorporate multiple channels of contagion. For example fire sales have been incorporated when the process does not have growth in Chapter 3. Incorporating firesales or other indirect contagion mechanisms adds another fixed point to the analysis, driven for example by the price dynamics of an illiquid asset. In the context with growth, the entire path of the endogenous price process would have to be consistent with the time threshold functions (4.7), which in turn are price dependent.

Finally, key research questions revolve around the optimal dividend distributions. In the Cramér Lundberg setting, this problem has been studied extensively in the literature, see e.g., [42, 128, 188]. For spectrally negative Lévy processes, the optimal dividend distribution have been shown to be of constant barrier type, see e.g., [40, 169]. The study of optimal dividend distributions in the presence of network risk remains an open problem. One may replace ruin time by a ruin time observed with Poissonian frequency, see for example [4, 5], which is equivalent to a Parisian ruin time with exponential grace period. Instead of dividend distribution towards outside entities, an alternative setup would be that of bail-ins. In such settings, agents would divert part of their growth to other agents with the goal of preventing their ruin. The optimal rate of contribution towards bailing in other nodes in the networks has been treated in the simple case of one firm and one subsidiary in [37]. In the multi-dimensional setup one can consider a central node that covers the shortfall of the other nodes with the pool of capital to which all agents contribute, these are "bail-ins" (a term coined in 2010 as the opposite of bailouts). The bail-in amounts are the necessary funds to “reflect" the risk processes back to an optimal positive level. Consequently, the central node risk process is a marked point process in which the inter-arrival times are the first passage times of the members capital process, and the jump sizes are the bail-in amounts. This process has thus a non-trivial dependence structure between interarrival times and jump sizes, that we leave for future research.

Part II

GRAPHON MEAN FIELD INTERACTING SYSTEMS

Chapter 5

Graphon Mean Field Backward Stochastic Differential Equations and Associated Dynamic Risk Measures

This chapter is based on paper [4] in the publication list of Section 1.5.

Abstract. We study graphon mean-field backward stochastic differential equations (BSDEs) with jumps and associated dynamic risk measures. We establish the existence, uniqueness and measurability of solutions under some regularity assumptions and provide some estimates for the solutions. We moreover prove the stability with respect to an interacting graphon particle systems, and obtain the convergence of an interacting mean-field particle system with inhomogeneous interactions to the graphon mean-field BSDE. We then provide some comparison theorems for the graphon mean-field BSDEs. As an application, we introduce the graphon dynamic risk measure induced by the solution of a graphon mean-field BSDE system and study its properties. We finally provide a dual representation theorem for the graphon dynamic risk measure in the convex case.

Keywords: Graphon mean field, BSDEs with jumps, inhomogeneous interacting system, dynamic risk measures.

5.1 Introduction

The study of mean-field systems with homogeneous interaction goes back to Boltzmann, Vlasov, McKean and others (see e.g. [33, 154, 172]). Backward Stochastic Differential Equations (BSDEs) of mean-field type have been early studied in [72, 73] and since then, nonlinear mean-field BSDEs with jumps have been intensively investigated, see e.g., [137, 167, 168]. Moreover, the theory of mean-field games, introduced by Lasry and Lions in [163] and Huang, Caines and Malhamé [141, 142], has raised a lot of attention these last years; see in particular the recent book [83] and references therein. Motivated by applications in various domains, mean-field systems and mean-field games on large networks have been explored for different random graph models, including Erdös-Rényi graph [102] and inhomogeneous random graphs [181]. Recently, the use of graphons has emerged in order to analyze heterogeneous interaction in mean-field systems and game theory, see in particular [47, 78, 79, 82]. Graphons have been developed by Lovász et al., see e.g. [67, 68, 170], as a natural continuum limit object for large dense graphs. Essentially, a graphon is a symmetric measurable function $G : I^2 \to I$, with $I := [0, 1]$ indexing a continuum of possible positions for nodes in the graph and $G(u, v)$ representing the edge density between nodes placed at u and v.

Bayraktar et al. consider in [47] heterogeneously interacting diffusive particle systems and their large population limit, which is a forward graphon SDE system. In [55], Bayraktar et al. study the forward-backward SDEs with graphon interactions depending on the forward component, and the propagation of chaos of the corresponding interacting particle system. In [162], Lacker and Soret study stochastic graphon games and use the graphon equilibrium to approximate the Nash equilibrium for corresponding large finite games on any graph which converges in cut norm.

In this chapter, we are interested in the general study of graphon mean-field BSDEs with jumps, and their associated dynamic risk measures, defined, similarly as in the classic case, (see e.g. [45, 89, 123, 183]) as the opposite of solutions of graphon mean-field BSDEs with jumps.

We extend [89], which studies the mean-field BSDE with jumps, by introducing the graphon interaction in the drift to capture the heterogeneous interactions. We establish the existence, uniqueness and measurability of solutions and prove the stability with respect to the interacting graphons. We prove convergence results of finite interacting particle systems to graphon mean-field BSDEs. Compared to [55] where the interactions are described by the forward components, our graphon system is fully backward coupled with jumps.

The chapter is organized as follows. In Section 5.2, we introduce the notation and the definition of graphon mean-field BSDEs with jumps. We establish the existence, uniqueness and measurability of the solution, and provide comparison theorems under a monotone condition. We also study the continuity of solution on the label index and the stability with respect to different graphons. In Section 5.3, we show the convergence result for an interacting mean-field particle system with heterogeneous interactions to the graphon mean-field BSDEs with jumps. Section 5.4 concentrates on the associated graphon dynamic risk measures and its properties. In Section 5.4.2, we provide a dual representation theorem for graphon mean-field BSDEs in the convex case. Section 5.5 concludes.

5.2 Graphon mean-field BSDEs with jumps

5.2.1 Notation and definitions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $I = [0,1]$ and $\{W_u : u \in I\}$ be a family of i.i.d. m-dimensional Brownian motions defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\{N_u(dt, de) : u \in I\}$ be a family of independent Poisson measures defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with compensator $\nu_u(de)dt$ such that ν_u is a σ-finite measure on $E := \mathbb{R}^m_*$, with $\mathbb{R}_* := \mathbb{R}\backslash\{0\}$, equipped with its Borelian σ-algebra $\mathcal{B}(E)$, for each $u \in I$. Let $\{\widetilde{N}_u(dt, de) : u \in I\}$ be their compensator processes. Let $\mathbb{F} = \{\mathcal{F}_t, t \geqslant 0\}$ be the natural filtration associated with $\{W_u : u \in I\}$ and $\{N_u(dt, de) : u \in I\}$.

Let $T > 0$ be a fixed time horizon. Denote by P the predictable σ-algebra on $[0,T] \times \Omega$.

Given a Polish space $\mathcal{S}$, denote by $\mathcal{D}([0,T], \mathcal{S})$ the space of RCLL (right continuous with left limits) functions from $[0,T]$ to $\mathcal{S}$, equipped with the Skorokhod topology. Let $\mathcal{D}_m := \mathcal{D}([0,T], \mathbb{R}^m)$. Denote by $\mathcal{P}(\mathcal{S})$ the space of probability measure on $\mathcal{S}$. For a random variable X, $\mathcal{L}(X)$ denotes the law of X.

We use the following notation.

- $L^2(\mathcal{F}_t)$ is the set of all $\mathcal{F}_t$-measurable and square integrable random variables, for $t \in [0,T]$.
- $\mathbb{H}^2$ is the set of real-valued predictable processes ϕ such that

$$\|\phi\|_{\mathbb{H}^2} := (\mathbb{E}[\int_0^T \phi_t^2 dt])^{1/2} < \infty.$$

- $L^2_{\nu_u}$ (for each $u \in I$) is the set of all measurable functions $\ell : E \mapsto \mathbb{R}$ such that

$$\|\ell\|_{\nu_u} := (\int_E |\ell(y)|^2 \nu_u(dy))^{1/2} < \infty.$$

  Note that $L^2_{\nu_u}$ is a Hilbert space equipped with the scalar product

$$\langle \ell_1, \ell_2 \rangle_{\nu_u} := \int_E \ell_1(y)\ell_2(y)\nu_u(dy).$$

- $\mathbb{H}^2_{\nu_u}$ (for each $u \in I$) is the set of all predictable processes ℓ such that

$$\|\ell\|_{\mathbb{H}^2_{\nu_u}} := (\mathbb{E}[\int_0^T \|\ell_t\|^2_{\nu_u} dt])^{1/2} < \infty.$$

- $\mathbb{S}^2$ is the set of real-valued RCLL adapted processes ϕ with

$$\|\phi\|_{\mathbb{S}^2} := (\mathbb{E}[\sup_{t \in [0,T]} |\phi_t|^2])^{1/2} < \infty.$$

- $\mathcal{M}\mathbb{H}^2$ is the set of all measurable functions X from I to $\mathbb{H}^2$: $u \mapsto X_u$, satisfying

$$\sup_{u \in I} \|X_u\|^2_{\mathbb{H}^2} = \sup_{u \in I} \mathbb{E}[\int_0^T |X_u(t)|^2 dt] < \infty.$$

  For $X \in \mathcal{M}\mathbb{H}^2$, we define the norm

$$\|X\|^I_{\mathbb{H}^2} := \sup_{u \in I} (\mathbb{E}[\int_0^T |X_u(t)|^2 dt])^{1/2}.$$

  We define $\mathcal{M}L^2(\mathcal{F}_t)$ and $\mathcal{M}\mathbb{S}^2$ similarly.

- $\mathcal{M}\mathbb{H}^2_\nu := (\mathbb{H}^2_{\nu_u})^{\otimes I}$ is the set of all families $\ell := \{\ell_u\}_{u \in I}$ such that

$$\sup_{u \in I} (\mathbb{E}[\int_0^T \|\ell_{u,t}\|^2_{\nu_u} dt])^{1/2} < \infty.$$

  For $X \in \mathcal{M}\mathbb{S}^2$, we define the norm

$$\|X\|^I_{\mathbb{S}^2} := \sup_{u \in I} (\mathbb{E}[\sup_{t \in [0,T]} |X_u(t)|^2])^{1/2},$$

  and, for $\ell \in \mathcal{M}\mathbb{H}^2_\nu$,

$$\|X\|^I_{\mathbb{H}^2_\nu} := \sup_{u \in I} (\mathbb{E}[\int_0^T \|\ell_{u,t}\|^2_{\nu_u} dt])^{1/2}.$$

  Sometimes we denote $\ell_u := (\ell_{u,t})_{t \geqslant 0}$.

- $L^{2,I}(\mathcal{F}_t)$ (for $t \in [0,T]$) is the space of all $\mathcal{F}_t$-measurable family of random variables $X := \{X_u\}_{u \in I}$ satisfying

$$\|X\|_{L^{2,I}} := (\mathbb{E}\Big[\int_I |X_u|^2 du\Big])^{1/2} < \infty.$$

  We define further the scalar product

$$\langle X, Y \rangle_{L^{2,I}} := \mathbb{E}[\int_I X_u Y_u du].$$

5.2.2 Graphons

A graphon is defined as a symmetric measurable function $G : I \times I \to I$, with $I = [0,1]$. Graphons can be regarded as the limits of edge matrices of weighted graphs, when the size of the graph (number of vertices) goes to infinity. Indeed, by relabelling vertices of the graph by i/n, $i \in [n] := \{1, \ldots, n\}$, as n becomes large, the labels $i/n, i \in [n]$ become close to each other, tending to a continuum in $[0,1]$. Let $\mathcal{B}(I)$ be the Borel algebra on I. The so-called cut norm of a graphon is defined by

$$\|G\|_\square := \sup_{A,B \in \mathcal{B}(I)} \Big|\int_{A \times B} G(u,v) du dv\Big|.$$

We can also view a graphon as an operator from $L^\infty(I)$ to $L^1(I)$, defined for any $\phi \in L^\infty(I)$ as:

$$G\phi(u) := \int_I G(u,v)\phi(v)dv.$$

By Lovász [170, Lemma 8.11], the resulting operator norm turns out to be equivalent to the cut norm:

$$\|G\|_\square \leqslant \|G\|_{\infty\to 1} \leqslant 4\|G\|_\square,$$

with

$$\|G\|_{\infty\to 1} := \sup_{|\phi|\leqslant 1} \|G\phi\|_{L^1}.$$

These norms will be used to study convergence issues when the graphon system is induced by a sequence of graphons. To get stronger convergence results, we shall need to introduce another operator norm for graphons, and consider G as an operator from $L^\infty(I)$ to $L^\infty(I)$, with the norm

$$\|G\|_{\infty\to\infty} := \sup_{|\phi|\leqslant 1} \|G\phi\|_{L^\infty}.$$

We are now ready to introduce the graphon mean-field BSDE with jumps:

$$\begin{aligned} X_u(t) = \xi_u + \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx)dyds - \int_t^T Z_u(s)dW_u(s) \\ - \int_t^T \int_E \ell_{u,s}(e)\tilde{N}_u(ds,de), \qquad u \in I, \quad \text{for} \quad t \in [0,T], \end{aligned} \tag{5.1}$$

where $\mu_y := \mathcal{L}(X_y) \in \mathcal{P}(\mathcal{D}_m)$ and $\mu_{y,s} := \mathcal{L}(X_y(s)) \in \mathcal{P}(\mathbb{R}^m)$. We assume that $\xi := \{\xi_u\}_{u\in I} \in \mathcal{M}L^2(\mathcal{F}_T)$, that is for each $u \in I$, $\xi_u \in L^2(\mathcal{F}_T)$ and the map $u \mapsto \xi_u$ is measurable.

We define the space

$$\begin{aligned} \mathcal{M} := \{\Phi := \{(X_u, Z_u, \ell_u(\cdot)) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{H}^2_{\nu_u}\}_{u\in I}, \text{ such that} \\ \|\Phi\|_{\mathcal{M}} := \sup_{u\in I}\big(\mathbb{E}[\sup_{t\in[0,T]} |X_u(t)|^2] + \mathbb{E}[\int_0^T |Z_u(t)|^2 dt + \mathbb{E}[\int_0^T \|\ell_{u,t}\|^2_{\nu_u} dt]\big)^{1/2} < \infty\}. \end{aligned}$$

We consider the following Wasserstein distances between two probability measures μ and ν:

$$\mathcal{W}_2(\mu,\nu) := (\inf\{\mathbb{E}[|X_1 - X_2|^2] : \mathcal{L}(X_1) = \mu, \mathcal{L}(X_2) = \nu\})^{1/2}, \quad \text{for} \quad \mu,\nu \in \mathcal{P}(\mathbb{R}^m),$$

$$\mathcal{W}_{2,T}(\mu,\nu) := (\inf\{\sup_{t\in[0,T]} \mathbb{E}|X_1(t) - X_2(t)|^2 : \mathcal{L}(X_1) = \mu, \mathcal{L}(X_2) = \nu\})^{1/2}, \quad \text{for} \quad \mu,\nu \in \mathcal{P}(\mathcal{D}_m).$$

Further, for two families of probability measures $\mu = \{\mu_u\}_{u\in I}$ and $\nu = \{\nu_u\}_{u\in I}$, we set

$$\mathcal{W}_2^{\mathcal{M}}(\mu,\nu) := \sup_{u\in I} \mathcal{W}_2(\mu_u,\nu_u), \text{ for } \mu,\nu \in \mathcal{P}(\mathcal{M}L^2(\mathcal{F}_t)) \text{ for all } t \in [0,T],$$

and

$$\mathcal{W}_{2,T}^{\mathcal{M}}(\mu,\nu) := \sup_{u\in I} W_{2,T}(\mu_u,\nu_u), \text{ for } \mu,\nu \in \mathcal{P}(\mathcal{M}\mathbb{S}^2).$$

Remark 5.1. *For each fixed $T > 0$, we have*

$$\mathcal{W}_2(\mu,\nu) \geqslant \sup_f \Big| \int_{\mathbb{R}^m} f(x)\mu(dx) - \int_{\mathbb{R}^m} f(x)\nu(dx) \Big|, \quad \mu,\nu \in \mathcal{P}(\mathbb{R}^m),$$

$$\mathcal{W}_{2,T}(\mu,\nu) \geqslant \sup_f \Big| \int_{\mathbb{R}^m} f(x)\mu_T(dx) - \int_{\mathbb{R}^m} f(x)\nu_T(dx) \Big|, \quad \mu,\nu \in \mathcal{P}(\mathcal{D}_m),$$

and

$$\mathcal{W}^{\mathcal{M}}_{2,T}(\mu,\nu)) \geqslant \sup_{u\in I}\sup_f \Big| \int_{\mathbb{R}^m} f(x)\mu_{u,T}(dx) - \int_{\mathbb{R}^m} f(x)\nu_{u,T}(dx) \Big|, \quad \mu,\nu \in \mathcal{M}\mathbb{S}^2,$$

where the supremum over f is taken over all Lipschitz continuous functions $f : \mathbb{R}^m \to \mathbb{R}$ with Lipschitz constant 1 such that the integral exists.

For notation simplicity, we restrict ourselves to the case $m = 1$. The proofs can be easily generalized to $m > 1$.

5.2.3 Existence and uniqueness results

In this section, we prove existence and uniqueness of solutions to the graphon mean-field BSDE system with jumps (5.1).

Definition 5.2. *A solution of the graphon mean-field BSDE system with jumps (5.1) consists of a family of processes $\Phi := (X_u, Z_u, \ell_u)_{u\in I}$ with $(X_u, Z_u, \ell_u) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{H}^2_{\nu_u}$ for all u in I satisfying (5.1), where X_u is a RCLL $\mathbb{R}$-valued optional process, and Z_u (resp. ℓ_u) is a $\mathbb{R}$-valued predictable process defined on $\Omega \times [0,T]$ (resp. $\Omega \times [0,T] \times E$) such that the stochastic integral is well defined.*

Canonical coupling. Note that in the graphon system, the solution for each label u can be influenced only by the law of the first component of the solution for other labels. Thus, when we couple the Brownian motions and Poisson random measures in (5.1), the law of the state for each label, $\mathcal{L}(X_u)$, remains unchanged, as proved in [55, Lemma 2.1] for the coupled graphon FBSDE system. To study the solution of the graphon BSDE system, we must require some form of measurability for $u \mapsto X_u$ such as weak-sense measurability for the law of $\mathcal{L}(X_u)$. However, through a suitable coupling, we can achieve strong measurability for X in the space $\mathcal{M}\mathbb{S}^2$ and transform the original graphon system into a fully coupled system defined in the canonical space. We refer to this as the canonical coupling, which simplifies some convergence analysis. We make the following assumption:

Assumption 5.1 (Intensity measure)**.** *For each $\omega \in [1,2]$, the function $I \ni u \mapsto \Phi_u^{-1}(\omega - 1) \in \mathbb{R}$, is measurable, where Φ_u denotes the cumulative distribution function of ν_u. We define $\Phi_u^{-1}(1)$ as the essential supremum and $\Phi_u^{-1}(0)$ as the essential infimum.*

Define the canonical filtered probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{F}}, \bar{\mathbb{P}})$, where $\bar{\mathbb{F}} = \{\bar{\mathcal{F}}_t, t \geqslant 0\}$ is the completed natural filtration and $\bar{\mathbb{P}}$ is the probability measure, both generated by a canonical one-dimensional Brownian motion W and a Poisson random measure $N(dt, de)$ with compensator $\nu(de)dt$. Here, the measure ν is uniform on $[1,2]$, and is called the canonical measure. The idea is to use a common

Poisson random measure to generate different random measures N_u for all $u \in I$ through the mapping $\Phi_u^{-1}, u \in I$. Note that Φ_u^{-1} is monotone and deterministic, it preserve the properties of Poisson random measure. $N(dt, \Phi_u^{-1}(e-1)de)$ is a Poisson random measure with intensity $\nu_u(de)dt$. The canonically coupled graphon system is now written as follows:

$$\begin{aligned}\bar{X}_u(t) = \xi_u + \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f(s, x, \bar{X}_u(s), \bar{Z}_u(s), \bar{\ell}_{u,s}(\cdot)) \mu_{y,s}(dx) dy ds - \int_t^T \bar{Z}_u(s) dW(s) \\ - \int_t^T \int_E \bar{\ell}_{u,s}(\Phi_u^{-1}(e-1)) \widetilde{N}(ds, de), \qquad u \in I, \quad \text{for} \quad t \in [0,T].\end{aligned} \tag{5.2}$$

Note that $\mathcal{L}(\bar{X}, \bar{Z}, \bar{\ell}) = \mathcal{L}(X, Z, \ell)$. We now give the Lipschitz conditions on the driver f:

Assumption 5.2. *For each $u \in I$,*

$$\begin{aligned}&f : \Omega \times [0,T] \times \mathbb{R}^3 \times L^2_{\nu_u} \to \mathbb{R} \\ &(\omega, t, x', x, z, \ell(\cdot)) \mapsto f(\omega, t, x', x, z, \ell(\cdot))\end{aligned}$$

is $P \otimes \mathcal{B}(\mathbb{R}^3) \otimes \mathcal{B}(L^2_{\nu_u})$ measurable, and satisfies $f(\cdot, \cdot, 0, 0, 0, 0) \in \mathbb{H}^2$, and f is Lipschitz-continuous in (x', x, z, ℓ), i.e., there exists a constant $C \geqslant 0$ such that $dt \otimes d\mathbb{P}$-a.s., for each (x'_1, x_1, z_1, ℓ_1) and (x'_2, x_2, z_2, ℓ_2), we have

$$\begin{aligned}&\left| f(\omega, t, x'_1, x_1, z_1, \ell_1(\cdot)) - f(\omega, t, x'_2, x_2, z_2, \ell_2(\cdot)) \right| \\ \leqslant & C(|x'_1 - x'_2| + |x_1 - x_2| + |z_1 - z_2| + \|\ell_1 - \ell_2\|_{\nu_u}).\end{aligned}$$

To prove the existence and uniqueness, we need the following lemma.

Lemma 5.3. *For a given $\overline{x} \in \mathcal{M}\mathbb{S}^2$, let $\overline{\mu}_{y,s} := \mathcal{L}(\overline{x}_y(s))$, and suppose Assumption 5.1 and 5.2 are satisfied. Then there exists a unique solution $(X_u, Z_u, \ell_u)_{u \in I} \in \mathcal{M}$ to the following graphon BSDE with jumps,*

$$\begin{aligned}X_u(t) &= \overline{x}_u(T) + \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f(s, x, X_u(s), Z_u(s), \ell_{u,s}(\cdot)) \overline{\mu}_{y,s}(dx) dy ds \\ &\qquad - \int_t^T Z_u(s) dW(s) - \int_t^T \int_E \ell_{u,s}(\Phi_u^{-1}(e-1)) \widetilde{N}(ds, de), \quad t \in [0,T], \\ X_u(T) &= \overline{x}_u(T).\end{aligned}$$

Moreover, X belong to $\mathcal{M}\mathbb{S}^2$.

Proof. Define the following iterating equations (for $n \geqslant 1$):

$$\begin{aligned}X_u^n(t) = \xi_u + \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f(s, x, X_u^{n-1}(s), Z_u^{n-1}(s), \ell_{u,s}^{n-1}(\cdot)) \overline{\mu}_{y,s}(dx) dy ds \\ - \int_t^T Z_u^n(s) dW(s) - \int_t^T \int_E \ell_{u,s}^n(\Phi_u^{-1}(e-1)) \widetilde{N}(ds, de), \quad t \in [0,T],\end{aligned} \tag{5.3}$$

where for $n = 0$ we set $\Phi_u^0 = (0,0,0)$ for all $u \in I$. For each $n \geqslant 1$, the driver

$$\widehat{f}_u^n(s,\cdot) := \int_I \int_{\mathbb{R}} G(u,y) f(s,x,X_u^{n-1}(s),Z_u^{n-1}(s),\ell_{u,s}^{n-1}(\cdot))\overline{\mu}_{y,s}(dx)dy$$

does not depend on $\Phi^n := (X_u^n, Z_u^n, \ell_u^n)_{u\in I}$, for each $u \in I$. Thus X_u^n is given by

$$X_u^n(t) = \mathbb{E}[\overline{x}_u(T) + \int_t^T \widehat{f}_u^n(s,\cdot)ds|\mathcal{F}_t]. \tag{5.4}$$

By the martingale representation theorem for locally square integrable martingales, see e.g. [191], (Z_u^n, ℓ_u^n) is the unique pair of predictable processes satisfying

$$\mathbb{E}[\overline{x}_u(T) + \int_0^T \widehat{f}_u^n(s,\cdot)ds|\mathcal{F}_t] = \overline{x}_u(0) + \int_0^t Z_u^n(s)dW(s) + \int_0^t \int_E \ell_{u,s}^n(\Phi_u^{-1}(e-1))\widetilde{N}(ds,de).$$

By our assumptions, $u \mapsto \bar{x}_u$ is measurable, $u \mapsto \Phi_u^{-1}(\omega-1)$ is measurable and, for all $(u,s) \in I\times[0,T]$, $\widehat{f}_u^n(s,\cdot)$ is Lipschitz-continuous in x. Suppose $(u,s) \mapsto X_u^{n-1}(s) \in L^2(\Omega,\mathcal{F},\mathbb{P})$ is measurable. By using [55, Lemma A.3 and A.4], $I \ni u \mapsto \int_0^T \widehat{f}_u^n(s,\cdot)ds \in L^2(\mathcal{F}_T)$ is measurable. Note that by Jensen's inequality, for $Y \in L^2(\mathcal{F}_T)$,

$$Y \mapsto \mathbb{E}[Y|\mathcal{F}_t], \quad t \in [0,T],$$

is a contraction and therefore is continuous. Hence we have that for any $t \in [0,T]$, $u \mapsto X_u^n(t)$ is measurable. It follows by [55, Lemma A.2] that $u \mapsto X_u^n$ is measurable for $n \geqslant 1$. Clearly $u \mapsto X_u^0$ is measurable, thus for all $n \geqslant 1$, X_u^n is measurable in u. Then by classic existence and uniqueness results for BSDEs with jumps (cf. [183, Proposition A.2]), for each $u \in I$, Φ_u^n converges to some limit $\Phi_u \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{H}_u^2$, and since $u \mapsto X_u^n$ is measurable, the limit $u \mapsto X_u$ is measurable and thus $X \in \mathcal{M}\mathbb{S}^2$. ☐

We are now ready to prove the following existence and uniqueness theorem.

Theorem 5.4. *We suppose Assumption 5.1 and 5.2 are satisfied and $\xi \in \mathcal{M}L^2(\mathcal{F}_T)$. Then the coupled system (5.2) admits a unique solution $\bar{\Phi} := (\bar{X},\bar{Z},\bar{\ell}) \in \mathcal{M}$ such that $\bar{X} \in \mathcal{M}\mathbb{S}^2$. Furthermore, the graphon mean-field BSDE system with jumps (5.1) also admits a unique solution $\Phi := (X,Z,\ell) \in \mathcal{M}$, and $I \ni u \mapsto \mathcal{L}(X_u)$ is measurable.*

Proof. The measurability issue has been addressed in Lemma 5.3. We prove the existence and uniqueness results by fixed point theorem arguments. With a slight abuse of notation, we couple the system by the canonic way introduced before without changing the form of (5.1). We establish a Picard iteration sequence and proceed by proving the contraction property. We first define the following iterating equation:

$$\begin{aligned} X_u^n(t) = \xi_u &+ \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f(s,x,X_u^n(s),Z_u^n(s),\ell_{u,s}^n(\cdot))\mu_{y,s}^{n-1}(dx)dyds \\ &- \int_t^T Z_u^n(s)dW_u(s) - \int_t^T \int_E \ell_{u,s}^n(e)\widetilde{N}_u(ds,de), \quad t \in [0,T], \end{aligned} \tag{5.5}$$

where $\mu_y^n := \mathcal{L}(X_y^n)$, and for $n = 0$ we set $\Phi_u^0 = (0,0,0)$ for all $u \in I$. By Lemma 5.3, there exists a unique solution to the above equation, for each $n \geqslant 1$. Let $\Phi^n := (X^n, Z^n, \ell^n(\cdot))$ be the solution of the above iterating BSDE with jumps with $X^n \in \mathcal{M}\mathbb{S}^2$ for $n \geqslant 1$. We define the mapping $\Psi(\Phi^{n-1}) = \Phi^n$.

We now show that Ψ is a contraction in $\mathcal{M}$. Let $\overline{X}_u^n(t) := X_u^n(t) - X_u^{n-1}(t)$, $\overline{Z}_u^n(t) := Z_u^n(t) - Z_u^{n-1}(t)$ and $\overline{\ell}_{u,t}^n := \ell_{u,t}^n - \ell_{u,t}^{n-1}$. For $r > 0$, applying Itô's formula to $e^{rs}|\overline{X}_u^n(s)|^2$ between 0 and T, $n \geqslant 1$, we obtain

$$
\begin{aligned}
&(\overline{X}_u^n(0))^2 + r\int_0^T e^{rs}(\overline{X}_u^n(s))^2 ds + \int_0^T e^{rs}(\overline{Z}_u^n(s))^2 ds + \int_0^T e^{rs}\|\overline{\ell}_{u,s}^n\|_u^2 ds \\
&= 2\int_0^T e^{rs}\overline{X}_u^n(s)\int_I \Big(\int_{\mathbb{R}} G(u,y) f(s,x,X_u^n(s),Z_u^n(s),\ell_{u,s}^n(\cdot))\mu_{y,s}^{n-1}(dx) \\
&\quad - \int_{\mathbb{R}} G(u,y) f(s,x,X_u^{n-1}(s),Z_u^{n-1}(s),\ell_{u,s}^{n-1}(\cdot))\mu_{y,s}^{n-2}(dx)\Big) dy ds \\
&\quad - 2\int_0^T e^{rs}\overline{X}_u^n(s)\overline{Z}_u^n(s) dW_u(s) - 2\int_0^T e^{rs}\int_E \overline{X}_u^n(s-)\overline{\ell}_{u,s}^n(e)\tilde{N}_u(ds,de).
\end{aligned}
$$

Taking expectation, noting that X^n and X^{n-1} belong to $\mathbb{S}^2$, all local martingales of the right hand side in the above inequality are martingales. We thus get

$$
\begin{aligned}
&\mathbb{E}\Big[r\int_0^T e^{rs}(\overline{X}_u^n(s))^2 ds + \int_0^T e^{rs}(\overline{Z}_u^n(s))^2 ds + \int_0^T e^{rs}\|\overline{\ell}_s^n\|_{\nu_u}^2 ds\Big] \\
&\leqslant \mathbb{E}\Big[2\int_0^T e^{rs}\overline{X}_u^n(s)\int_I \Big(\int_{\mathbb{R}} G(u,y) f(s,x,X_u^n(s),Z_u^n(s),\ell_{u,s}^n(\cdot))\mu_{y,s}^{n-1}(dx) \\
&\quad - \int_{\mathbb{R}} G(u,y) f(s,x,X_u^{n-1}(s),Z_u^{n-1}(s),\ell_{u,s}^{n-1}(\cdot))\mu_{y,s}^{n-2}(dx)\Big) dy ds\Big].
\end{aligned}
$$

Let

$$
\begin{aligned}
A := &\int_{\mathbb{R}} G(u,y) f(s,x,X_u^n(s),Z_u^n(s),\ell_{u,s}^n(\cdot))\mu_{y,s}^{n-1}(dx) \\
&- \int_{\mathbb{R}} G(u,y) f(s,x,X_u^{n-1}(s),Z_u^{n-1}(s),\ell_{u,s}^{n-1}(\cdot))\mu_{y,s}^{n-2}(dx).
\end{aligned}
$$

The Lipschitz property of the driver f and boundedness of the graphon G imply that for some constant C_0, we have

$$
\begin{aligned}
A^2 \leqslant &2\Big|\int_{\mathbb{R}} G(u,y)[f(s,x,\Phi_u^n(s)) - f(s,x,\Phi_u^{n-1}(s))]\mu_{y,s}^{n-1}(dx)\Big|^2 \\
&+ 2\Big|\int_{\mathbb{R}} G(u,y) f(s,x,\Phi_u^{n-1}(s))[\mu_{y,s}^{n-1} - \mu_{y,s}^{n-2}](dx)\Big|^2 \\
\leqslant &C_0(|\overline{X}_u^n(s)|^2 + |\overline{Z}_u^n(s)|^2 + \|\overline{\ell}_{u,s}^n\|_{\nu_u}^2) + C_0(\mathcal{W}_2(\mu_{y,s}^{n-1},\mu_{y,s}^{n-2}))^2,
\end{aligned}
$$

where the last inequality uses Remark 5.1. Then for any $\varepsilon > 0$, by using the inequality $2ab \leqslant$

$a^2/\varepsilon^2 + \varepsilon^2 b^2$, we have

$$
\begin{aligned}
&\mathbb{E}[r\int_0^T e^{rs}(\overline{X}_u^n(s))^2 ds + \int_0^T e^{rs}(\overline{Z}_u^n(s))^2 ds + \int_0^T e^{rs}\|\overline{\ell}_s^n\|_u^2 ds] \\
\leqslant& C_0^2\varepsilon^2\mathbb{E}[\int_0^T e^{rs}(|\overline{X}_u^n(s)|^2 + |\overline{Z}_u^n(s)|^2 + \|\overline{\ell}_{u,s}^n\|_{\nu_u}^2 + \int_I (\mathcal{W}_2(\mu_{y,s}^{n-1}, \mu_{y,s}^{n-2}))^2 dy) ds] \\
&+ \frac{1}{\varepsilon^2}\mathbb{E}[\int_0^T e^{rs}|\overline{X}_u^n(s)|^2 ds] \\
\leqslant& C_0^2\varepsilon^2\mathbb{E}[\int_0^T e^{rs}(|\overline{X}_u^n(s)|^2 + |\overline{Z}_u^n(s)|^2 + \|\overline{\ell}_{u,s}^n\|_{\nu_u}^2) ds + \int_I\int_0^T e^{rs}(\mathcal{W}_2(\mu_{y,s}^{n-1}, \mu_{y,s}^{n-2}))^2 ds dy] \\
&+ \frac{1}{\varepsilon^2}\mathbb{E}[\int_0^T e^{rs}|\overline{X}_u^n(s)|^2 ds] \\
\leqslant& C_0^2\varepsilon^2\{\mathbb{E}[\int_0^T e^{rs}(|\overline{X}_u^n(s)|^2 + |\overline{Z}_u^n(s)|^2 + \|\overline{\ell}_{u,s}^n\|_{\nu_u}^2) ds] + \sup_{y\in I}\mathbb{E}[\int_0^T e^{rs}|\overline{X}_y^{n-1}(s)|^2 ds]\} \\
&+ \frac{1}{\varepsilon^2}\mathbb{E}[\int_0^T e^{rs}|\overline{X}_u^n(s)|^2 ds].
\end{aligned}
$$

By choosing appropriate r and ε such that $r - C_0^2\varepsilon^2 - \frac{1}{\varepsilon^2} > C_0^2\varepsilon^2$ and $1 - C_0^2\varepsilon^2 > 0$, we obtain the contraction inequality with certain constant $\alpha < 1$:

$$
\sup_{u\in I}\mathbb{E}[\int_0^T e^{rs}(|\overline{X}^n|^2 + |\overline{Z}^n|^2 + \|\overline{\ell}_u^n\|_{\nu_u}^2) ds] \leqslant \alpha \sup_{u\in I}\mathbb{E}[\int_0^T e^{rs}(|\overline{X}^{n-1}|^2 + |\overline{Z}^{n-1}|^2 + \|\overline{\ell}_u^{n-1}\|_{\nu_u}^2) ds].
$$

We further get the contraction inequality in $\mathcal{M}$,

$$
\sup_{u\in I}\mathbb{E}[\int_0^T (|\overline{X}^n|^2 + |\overline{Z}^n|^2 + \|\overline{\ell}_u^n\|_{\nu_u}^2) ds] \leqslant \alpha \sup_{u\in I}\mathbb{E}[\int_0^T (|\overline{X}^{n-1}|^2 + |\overline{Z}^{n-1}|^2 + \|\overline{\ell}_u^{n-1}\|_{\nu_u}^2) ds],
$$

which implies that the map Ψ is a contraction in $\mathcal{M}$. It thus has a unique fixed point, denoted as $\Phi := (X, Z, \ell)$. Now taking the limit in the iterating equation (5.3), we conclude that Φ is the unique solution of (5.1). Since $u \mapsto X_u^n$ is measurable for each $n \geqslant 1$, the limit $u \mapsto X_u$ is also measurable in u.

We have proved the existence and uniqueness of a solution for the coupled system (5.2). The existence and uniqueness of a solution with the same law of the first component for the original graphon system (5.1) follows. The measurability for the map $u \mapsto \mathcal{L}(X_u)$ is a direct consequence since the weak topology is weaker than the topology induced by the running supremum of square expectation. The proof is then complete. □

Let $\mathbb{E}_t$ denote the conditional expectation given $\mathcal{F}_t$. We have the following estimate for the solution of the graphon mean-field BSDE .

Proposition 5.5. *Suppose Assumption 5.2 holds and let (X, Z, ℓ) be the solution of the graphon mean-field BSDE (5.1) with terminal value $\xi \in \mathcal{M}L^2(\mathcal{F}_T)$. Denote by C_L the Lipschitz constant of f. Let*

$\eta, r > 0$ be some constants such that $\eta < 1/(4C_L^2)$. If $r - 2/\eta \geqslant 2$, we have

$$\int_I \mathbb{E}[e^{rt}|X_u(t)|]du \leqslant \int_I \mathbb{E}[e^{rT}|\xi_u|^2]du + 4\eta \int_t^T e^{rs} \int_I G^2(u,y) f^2(s,0,0,0,0)dyds,$$

and for each $u \in I$,

$$\mathbb{E}[e^{rt}|X_u(t)|^2] \leqslant \big(1 + 4\eta C_L^2(T-t)\big)\big(4\eta \int_t^T e^{rs} \int_I G^2(u,y) f^2(s,0,0,0,0)dyds\big) + 4\eta C_L^2(T-t) \int_I \mathbb{E}[e^{rT}|\xi_u|^2]du + \mathbb{E}[e^{rT}|\xi_u|^2].$$

Furthermore, we get the following estimate without the expectation:

$$e^{rt}|X_u(t)|^2 \leqslant \big(1 + 4\eta C_L^2(T-t)\big)\big(4\eta \int_t^T e^{rs} \int_I G^2(u,y) f^2(s,0,0,0,0)dyds\big) + 4\eta C_L^2(T-t) \int_I \mathbb{E}_t[e^{rT}|\xi_u|^2]du + \mathbb{E}_t[e^{rT}|\xi_u|^2].$$

Proof. For any $u \in I$, by applying Itô's formula to $e^{rs}|X_u(s)|^2$ between $[t,T]$ and taking conditional expectation given $\mathcal{F}_t$, we obtain

$$\begin{aligned} &e^{rt}|X_u(t)|^2 + r\mathbb{E}_t\big[\int_t^T e^{rs}|X_u(s)|^2 ds\big] + \mathbb{E}_t\big[\int_t^T e^{rs}|Z_u(s)|^2 ds\big] + \mathbb{E}_t\big[\int_t^T e^{rs}\|\ell_{u,t}\|_{\nu_u}^2 ds \\ &\leqslant 2\mathbb{E}_t\big[\int_t^T e^{rs} X_u(s) \int_I \int_{\mathbb{R}} G(u,y) f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx)dyds\big] + \mathbb{E}_t[e^{rT}|X_u(T)|^2]. \end{aligned}$$

By using the inequality $2ab \leqslant a^2/\varepsilon^2 + \varepsilon^2 b^2$ and the Lipschitz property of f, we have

$$
\begin{aligned}
& e^{rt}|X_u(t)|^2 + r\mathbb{E}_t\Big[\int_t^T e^{rs}|X_u(s)|^2 ds\Big] + \mathbb{E}_t\Big[\int_t^T e^{rs}|Z_u(s)|^2 ds\Big] + \mathbb{E}_t\Big[\int_t^T e^{rs}\|\ell_{u,t}\|^2_{\nu_u} ds\Big] \\
& \leqslant 2\mathbb{E}_t\Big[\frac{1}{\varepsilon^2}\int_t^T e^{rs}|X_u(s)|^2 ds + \varepsilon^2 \int_t^T e^{rs}\Big|\int_I\int_{\mathbb{R}} G(u,y) f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx)dy\Big|^2 ds\Big] \\
& + \mathbb{E}_t[e^{rT}|X_u(T)|^2] \\
& \leqslant 2\mathbb{E}_t\Big[\frac{1}{\varepsilon^2}\int_t^T e^{rs}|X_u(s)|^2 ds\Big] + \mathbb{E}_t[e^{rT}|X_u(T)|^2] \\
& + \varepsilon^2\mathbb{E}_t\Big[\int_t^T e^{rs}\Big|\int_I\int_{\mathbb{R}} G(u,y)|f(s,0,0,0,0) + C_L(|x| + |X_u(s)| + |Z_u(s)| + \|\ell_{u,s}\|_{\nu_u})|\mu_{y,s}(dx)dy\Big|^2 ds\Big] \\
& \leqslant 2\mathbb{E}_t\Big[\frac{1}{\varepsilon^2}\int_t^T e^{rs}|X_u(s)|^2 ds\Big] + \mathbb{E}_t[e^{rT}|X_u(T)|^2] \\
& + 5\varepsilon^2\mathbb{E}_t\Big[\int_t^T e^{rs}\int_I\int_{\mathbb{R}} G^2(u,y)\big(f^2(s,0,0,0,0) + C_L^2(|x|^2 + |X_u(s)|^2 + |Z_u(s)|^2 + \|\ell_{u,s}\|^2_{\nu_u})|\mu_{y,s}(dx)dyds\Big] \\
& \leqslant 2\mathbb{E}_t\Big[\frac{1}{\varepsilon^2}\int_t^T e^{rs}|X_u(s)|^2 ds\Big] + \mathbb{E}_t[e^{rT}|X_u(T)|^2] + 5\varepsilon^2\int_t^T e^{rs}\int_I G^2(u,y)f^2(s,0,0,0,0)dyds \\
& + 5\varepsilon^2 C_L^2\mathbb{E}_t\Big[\int_t^T e^{rs}(|X_u(s)|^2 + |Z_u(s)|^2 + \|\ell_{u,s}\|^2_{\nu_u}))ds\Big] + 5\varepsilon^2 C_L^2\int_t^T e^{rs}\int_I \mathbb{E}|X_y(s)|^2 dyds.
\end{aligned}
\tag{5.6}
$$

Taking expectation with respect to $\mathcal{F}$ and integrating for all $u \in I$, we get

$$
\begin{aligned}
& \int_I\{\mathbb{E}[e^{rt}|X_u(t)|^2] + r\mathbb{E}\Big[\int_t^T e^{rs}|X_u(s)|^2 ds\Big] + \mathbb{E}\Big[\int_t^T e^{rs}|Z_u(s)|^2 ds\Big] + \mathbb{E}\Big[\int_t^T e^{rs}\|\ell_{u,t}\|^2_{\nu_u} ds\}du \\
& \leqslant 2\int_I \mathbb{E}\Big[\frac{1}{\varepsilon^2}\int_t^T e^{rs}|X_u(s)|^2 ds\Big]du + \int_I \mathbb{E}[e^{rT}|X_u(T)|^2]du + 5\varepsilon^2\int_t^T e^{rs}\int_I G^2(u,y)f^2(s,0,0,0,0)dyds \\
& + 5\varepsilon^2 C_L^2\int_I \mathbb{E}\Big[\int_t^T e^{rs}(|X_u(s)|^2 + |Z_u(s)|^2 + \|\ell_{u,s}\|^2_{\nu_u}))ds\Big]du + 5\varepsilon^2 C_L^2\int_t^T e^{rs}\int_I \mathbb{E}|X_y(s)|^2 dyds.
\end{aligned}
\tag{5.7}
$$

Taking ε and r which satisfy $2/\varepsilon^2 + 10\varepsilon^2 C_L^2 < r$ and $10\varepsilon^2 C_L^2 < 1$ at the same time, i.e., satisfying the conditions in the proposition with change of variable $\eta = \varepsilon^2$, it follows that

$$
\int_I \mathbb{E}[e^{rt}|X_u(t)|^2]du \leqslant \int_I \mathbb{E}[e^{rT}|\xi_u|^2]du + 4\eta\int_t^T e^{rs}\int_I G^2(u,y)f^2(s,0,0,0,0)dyds. \tag{5.8}
$$

Now, inserting the above result into (5.6), we obtain for each $u \in I$,

$$
\begin{aligned}
\mathbb{E}[e^{rt}|X_u(t)|^2] \leqslant & \big(1 + 5\varepsilon^2 C_L^2(T-t)\big)\big(5\varepsilon^2\int_t^T e^{rs}\int_I G^2(u,y)f^2(s,0,0,0,0)dyds\big) \\
& + 5\varepsilon^2 C_L^2(T-t)\int_I \mathbb{E}[e^{rT}|\xi_u|^2]du + \mathbb{E}[e^{rT}|\xi_u|^2].
\end{aligned}
$$

Finally, by using equation (5.8) in the last line of (5.6), we obtain the last assertion. The proof is now complete. □

5.2.4 Comparison theorems

For convenience, let F_u denote the drift driver of the u component in the graphon mean-field system (5.1), i.e.

$$F_u(\omega, t, \mathcal{L}(X_t), x, z, \ell(\cdot)) := \int_I \int_{\mathbb{R}} G(u,y) f(t, x', x, z, \ell(\cdot)) \mu_{y,t}(dx') dy.$$

In order to compare the first components of the solutions of two graphon mean-filed BSDEs, we need the following additional assumption.

Assumption 5.3. *We assume that for each $u \in I$ and each $(x', x, z, \ell_1, \ell_2) \in \mathbb{R}^3 \times (L^2_{\nu_u})^2$, there exists a function $\phi^{x',x,z,\ell_1,\ell_2}_{u,t} \in L^2_{\nu_u}$ such that*

$$f(t, x', x, z, \ell_1) - f(t, x', x, z, \ell_2) \geqslant \langle \phi^{x',x,z,\ell_1,\ell_2}_{u,t}, \ell_1 - \ell_2 \rangle_{\nu_u},$$

with

$$\begin{aligned} \phi^{x',x,z,\ell_1,\ell_2}_{u,t} : \quad & [0,T] \times \Omega \times \mathbb{R}^3 \times (L^2_{\nu_u})^2 \mapsto L^2_{\nu_u}; \\ & (t, \omega, x', x, z, \ell_1, \ell_2) \mapsto \phi^{x',x,z,\ell_1,\ell_2}_{u,t}(\omega, \cdot) \end{aligned}$$

$P \otimes \mathcal{B}(\mathbb{R}^3) \otimes \mathcal{B}((L^2_{\nu_u})^2)$ measurable, bounded and satisfying $d\mathbb{P} \otimes dt \otimes d\nu_u$ a.s.

$$\phi^{x',x,z,\ell_1,\ell_2}_{u,t}(y) \geqslant -1 \quad \text{and} \quad |\phi^{x',x,z,\ell_1,\ell_2}_{u,t}(y)| \leqslant \psi(y),$$

for some $\psi \in L^2_{\nu_u}$.

We have the following comparison theorems.

Theorem 5.6 (Comparison theorem for graphon mean-field BSDE)**.** *Let $\xi^1, \xi^2 \in \mathcal{M}L^2(\mathcal{F}_T)$ and denote by (X^1, Z^1, ℓ^1) and (X^2, Z^2, ℓ^2) the solution of graphon mean-field BSDE with jumps (5.1) associated to (ξ^1, f_1) and (ξ^2, f_2) respectively. Let f_1 and f_2 both satisfy Assumption 5.2, and further assume that:*

- *At least one of f_1 and f_2 satisfies Assumption 5.3, and the other one (or at least one if both satisfy Assumption 5.3) is non-decreasing in x';*
- *For each $u \in I \backslash H$ with H a zero Lebesgue measure subset of I, $\xi^2_u \geqslant \xi^1_u$ a.s. and $f_2(\omega, t, x', x, z, \ell) \geqslant f_1(\omega, t, x', x, z, \ell)$ a.s. for all $(t, x', x, z, \ell) \in \mathbb{R}^4 \times L^2_{\nu_u}$.*

Then for all $t \in [0,T]$ and $u \in I \backslash H$, we have $X^2_u(t) \geqslant X^1_u(t)$ almost surely.

Proof. Without loss of generality, assume that f_1 satisfies Assumption 5.3 and f_2 is non-decreasing in x'. For each $u \in I$ and $i = 1, 2$, we denote by $(X_u^{i,n}, Z_u^{i,n}, \ell_u^{i,n})$ the solution of the following iterating BSDE with jumps,

$$\begin{aligned} X_u^{i,n}(t) = \xi_u^i + \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f_i(s, x, X_u^{i,n}(s), Z_u^{i,n}(s), \ell_{u,s}^{i,n}(\cdot)) \nu_{y,s}^{i,n-1}(dx) dy ds \\ - \int_t^T Z_u^{i,n}(s) dW_u^i(s) - \int_t^T \int_E \ell_{u,s}^{i,n}(e) \widetilde{N}_u^i(ds, de), \quad t \in [0,T], \end{aligned}$$

for $n \geqslant 1$, where $\nu_y^{i,n} := \mathcal{L}(X_y^{i,n})$, and for $n = 0$ we set $(X_u^{i,0}, Z_u^{i,0}, \ell_u^{i,0}) = (0,0,0)$ for all $u \in I$. For $i = 1, 2$, let

$$F_u^i(\omega, s, \mathcal{L}(X_s), x, z, \ell(\cdot)) := \int_I \int_{\mathbb{R}} G(u,y) f_i(s, x', x, z, \ell(\cdot)) \mu_{y,s}(dx') dy.$$

By our assumptions, we have

$$\begin{aligned} &\int_I \int_{\mathbb{R}} G(u,y) f_1(s, x, X_u^{1,0}(s), Z_u^{1,0}(s), \ell_{u,s}^{1,0}(\cdot)) \nu_{y,s}^{1,0}(dx) dy \\ &\quad \leqslant \int_I \int_{\mathbb{R}} G(u,y) f_2(s, x, X_u^{2,0}(s), Z_u^{2,0}(s), \ell_{u,s}^{2,0}(\cdot)) \nu_{y,s}^{2,0}(dx) dy. \end{aligned}$$

Moreover, since f_1 satisfies Assumption 5.3, the graphon mean-field driver $F_u^1(s, \mathcal{L}(X_s^{1,0}), x, z, \ell)$ satisfies the conditions in [183, Theorem 4.2], and from that comparison theorem for BSDE with jumps, we have for $t \in [0,T]$ and $u \in I \backslash H$,

$$X_u^{1,1}(t) \leqslant X_u^{2,1}(t) \qquad a.s.$$

Then, since f_2 is non-decreasing in x', we have for $u \in I \backslash H$,

$$F_u^1(s, \mathcal{L}(X_s^{1,1}), x, z, \ell) \leqslant F_u^2(s, \mathcal{L}(X_s^{1,1}), x, z, \ell) \leqslant F_u^2(s, \mathcal{L}(X_s^{2,1}), x, z, \ell).$$

Again by the classic comparison theorem, we obtain

$$X_u^{1,2}(t) \leqslant X_u^{2,2}(t) \qquad a.s.$$

Proceeding by the same argument as above, we iteratively obtain that, for $t \in [0,T]$ and $u \in I \backslash H$,

$$X_u^{1,n}(t) \leqslant X_u^{2,n}(t) \qquad a.s., \text{ for } \quad n \geqslant 1.$$

From the proof of the existence and uniqueness result, Theorem 5.4, we know that for $i = 1, 2$, $(X_u^{i,n}, Z_u^{i,n}, \ell_u^{i,n})$ converges to the unique solution associated to f_i respectively, denoted by (X^i, Z^i, ℓ^i). It thus follows that for all $u \in I \backslash H$ and $t \in [0,T]$, a.s. $X_u^1(t) \leqslant X_u^2(t)$, as desired. ☐

Theorem 5.7. *(Strict comparison for graphon mean-field BSDE) Suppose the assumptions in Theorem 5.6 hold. Further, assume that f_1 satisfies Assumption 5.3 with strict inequality, i.e., $\phi_{u,t}^{x',x,z,\ell_1,\ell_2}(y) > -1$, and $\xi_u^1 \geqslant \xi_u^2$ a.s. for each $u \in I \backslash H$ with H a zero Lebesgue measure subset of I, and $f_1(\omega, t, x', x, z, \ell) \geqslant f_2(\omega, t, x', x, z, \ell)$ a.s. for all $(t, x', x, z, \ell) \in \mathbb{R}^4 \times L_{\nu_u}^2$. Then if $X^1(t_0) = X^2(t_0)$ (i.e., $X_u^1(t_0) = X_u^2(t_0)$ for all $u \in I \backslash H$) for some $t_0 \in [0,T]$, we have $X^1(\cdot) = X^2(\cdot)$ a.s. on $[t_0, T]$, and $f_2(\omega, t, x', x, z, \ell) = f_1(\omega, t, x', x, z, \ell)$ on $[t_0, T]$ for $u \in I \backslash H$.*

Proof. For each $u \in I\backslash H$, let $\overline{X}_u(t) := X_u^1(t) - X_u^2(t)$, $\overline{Z}_u(t) := Z_u^1(t) - Z_u^2(t)$ and $\overline{\ell}_{u,t}(\cdot) := \ell_{u,t}^1(\cdot) - \ell_{u,t}^2(\cdot)$. Denote

$$\overline{F}_u(t) := F_u^1(t, \mathcal{L}(X_t^1), X_u^1, Z_u^1, \ell_{u,t}^1(\cdot)) - F_u^2(t, \mathcal{L}(X_t^2), X_u^2, Z_u^2, \ell_{u,t}^2(\cdot)).$$

First note that from Assumption 5.3, we have by adding and subtracting terms that

$$\begin{aligned}\overline{F}_u(t) \geqslant & F_u^1(t, \mathcal{L}(X_t^1), X_u^2, Z_u^2, \ell_{u,t}^2(\cdot)) - F_u^2(t, \mathcal{L}(X_t^2), X_u^2, Z_u^2, \ell_{u,t}^2(\cdot)) \\ & + \delta_{u,t}\overline{X}_u(t) + \pi_u(t)\overline{Z}_u(t) + \langle \phi_{u,t}, \overline{\ell}_{u,t}\rangle,\end{aligned}$$

where

$$\delta_{u,t} := \frac{F_u^1(t, \mathcal{L}(X_t^1), X_u^1, Z_u^1, \ell_{u,t}^1(\cdot)) - F_u^1(t, \mathcal{L}(X_t^1), X_u^2, Z_u^1, \ell_{u,t}^1(\cdot))}{\overline{X}_u(t)} \mathbb{1}_{\{\overline{X}_u(t) \neq 0\}},$$

$$\pi_{u,t} := \frac{F_u^1(t, \mathcal{L}(X_t^1), X_u^2, Z_u^1, \ell_{u,t}^1(\cdot)) - F_u^1(t, \mathcal{L}(X_t^1), X_u^2, Z_u^2, \ell_{u,t}^1(\cdot))}{\overline{X}_u(t)} \mathbb{1}_{\{\overline{Z}_u(t) \neq 0\}},$$

and $\phi_{u,t}$ is as given in Assumption 5.3, and is bounded. By the Lipschitz property of f on (x, z), δ_u and π_u are also bounded and predictable. For each $t_0 \in [0, T]$, let $(\Gamma_{t_0,t}^u)_{t \in [t_0,T]}$ be the unique solution of the SDE

$$d\Gamma_{t_0,t}^u = \Gamma_{t_0,t-}^u \big[\delta_{u,t}dt + \pi_{u,t}dW_u(t) + \int_E \phi_{u,t}(e)\widetilde{N}_u(dt, de)\big], \qquad \Gamma_{t_0,t_0}^u = 1.$$

Then by the classic comparison theorem for linear BSDE (cf. [183, Lemma 4.1]), we have for each $u \in I\backslash H$ and $t_0 \leqslant t \leqslant T$,

$$\overline{X}_u(t_0) \geqslant \mathbb{E}[\Gamma_{t_0,T}^u(\xi_u^1 - \xi_u^2) + \int_{t_0}^T \Gamma_{t_0,s}^u \overline{\phi}_{u,s} ds | \mathcal{F}_{t_0}], \tag{5.9}$$

where

$$\overline{\phi}_{u,t} := F_u^1(t, \mathcal{L}(X_t^1), X_u^2, Z_u^2, \ell_{u,t}^2(\cdot)) - F_u^2(t, \mathcal{L}(X_t^2), X_u^2, Z_u^2, \ell_{u,t}^2(\cdot)).$$

By Theorem 5.6, for each $u \in I\backslash H$, we have a.s. $X_u^1 \geqslant X_u^2$. By our assumptions and the non-decreasing property of one of the two functions f_1 and f_2, we have

$$\overline{\phi}_{u,t} = F_u^1(t, \mathcal{L}(X_t^1), X_u^2, Z_u^2, \ell_{u,t}^2(\cdot)) - F_u^2(t, \mathcal{L}(X_t^2), X_u^2, Z_u^2, \ell_{u,t}^2(\cdot)) \geqslant 0.$$

Since for each $u \in I\backslash H$, $\pi_{u,t}(e) > 1$ $d\mathbb{P} \otimes dt \otimes d\nu_u$-a.s., it follows by [183, Corollary 3.5] that $\Gamma_{t_0,t}^u > 0$ for all $t_0 \leqslant t \leqslant T$. Hence, by using Equation (5.9), we conclude that $X_u^1(t) = X_u^2(t)$ for all $u \in I\backslash H$, a.s. on $[t_0, T]$, and $f_2(\omega, t, x', x, z, \ell) = f_1(\omega, t, x', x, z, \ell)$ on $[t_0, T]$. □

5.2.5 Continuity and stability results

We study below the continuity and stability of our graphon mean-field BSDE system (5.1).

We need the following assumption:

Assumption 5.4. *For each $u \in I$,*

(i) $u \to \mathcal{L}(\xi_u)$ is continuous w.r.t. the $\mathcal{W}_2$ metric.

(ii) there exists a finite collection of intervals $\{I_i : i = 1, \dots, N\}$ such that $I = \cup_i I_i$, and for each $i \in \{1, \dots, N\}$, we have $G(u, v)$ is continuous at u for each $v \in I \backslash H_i$ for some zero Lebesgue measure set H_i.

For Lipschitz continuity, we need a stronger assumption.

Assumption 5.5. *There exists a finite collection of intervals $\{I_i : i = 1, \dots, N\}$ such that $I = \cup_i I_i$, and for some constant C, we have for all $u_1, u_2 \in I_i$, $v_1, v_2 \in I_j$, and $i, j \in \{1, \dots, N\}$,*

$$\mathcal{W}_2(\mathcal{L}(\xi_{u_1}), \mathcal{L}(\xi_{u_2})) \leqslant C|u_1 - u_2|,$$

and,

$$|G(u_1, v_1) - G(u_2, v_2)| \leqslant C(|u_1 - u_2| + |v_1 - v_2|).$$

To study the continuity of solutions with respect to the label u, we need to measure how close two solutions become in the usual norm for the solution of a BSDE with jumps as two labels $u_1, u_2 \in I$ get progressively closer. We will proceed by estimating the distance between two solutions through canonical coupling, which will then allow us to establish the continuity of solutions in the Wasserstein L^2 distance.

The following proposition gives the continuity and Lipschitz continuity of the graphon mean-field BSDE system (5.1).

Proposition 5.8. *Suppose that Assumption 5.2 holds and the measures $\{\nu_u\}_{u \in I}$ are a common measure ν. We have the following:*

(i) (Distance estimation) For each $i \in \{1, \dots, N\}$ and for all $u_1, u_2 \in I_i$, under the canonical coupling, we have (for some constant C)

$$\begin{aligned}
&\mathbb{E}[\sup_{t \in [0,T]} |X_{u_1}(t) - X_{u_2}(t)|^2] + \mathbb{E}[\int_0^T |Z_{u_1}(s) - Z_{u_2}(s)|^2 ds] + \mathbb{E}[\int_0^T \|\ell_{u_1,t} - \ell_{u_2,s}\|_\nu^2 ds] \\
&\quad \leqslant C\mathbb{E}|X_{u_1}(T) - X_{u_2}(T)|^2 + C\int_I |G(u_1, y) - G(u_2, y)|^2 dy.
\end{aligned}$$

(ii) (Continuity) Under Assumption 5.4, for each $i \in \{1, \dots, N\}$, the map $I_i \ni u \to \mathcal{L}(X_u)$ is continuous w.r.t. the $\mathcal{W}_{2,T}$ metric.

(iii) (Lipschitz continuity) Under Assumption 5.5, for each $i \in \{1, \dots, N\}$, the map $I_i \ni u \to \mathcal{L}(X_u)$ is Lipschitz continuous.

Proof. We omit some details since the proof is similar to that of Theorem 5.14. Fix $u_1, u_2 \in I$ and $t \in [0,T]$. Similar as in the proof of [47, Theorem 2.1], couple X_{u_1} and X_{u_2} with the same Brownian motion $W_{u_1} = W_{u_2} = W$ and the same Poisson measure $N_{u_1}(dt,de) = N_{u_2}(dt,de) = N(dt,de)$. By applying Itô's formula to $|X_{u_1}(t) - X_{u_2}(t)|$ and taking conditional expectation given $\mathcal{F}_t$, denoted by $\mathbb{E}_t$, we have

$$
\begin{aligned}
&\mathbb{E}_t|X_{u_1}(t) - X_{u_2}(t)|^2 + \mathbb{E}_t[\int_t^T |Z_{u_1}(s) - Z_{u_2}(s)|^2 ds] + \mathbb{E}_t[\int_t^T \|\ell_{u_1,s} - \ell_{u_2,s}\|_\nu^2 ds] \\
\leqslant& 2\mathbb{E}_t\Big[\int_t^T (X_{u_1}(s) - X_{u_2}(s)) \int_I \Big(\int_{\mathbb{R}} G(u_1,y) f(s,x,X_{u_1}(s),Z_{u_1}(s),\ell_{u_1,s}(\cdot))\mu_{y,s}(dx) \\
&\quad - \int_{\mathbb{R}} G(u_2,y) f(s,x,X_{u_2}(s),Z_{u_2}(s),\ell_{u_2,s}(\cdot))\mu_{y,s}(dx)\Big) dy ds\Big] + \mathbb{E}_t|X_{u_1}(T) - X_{u_2}(T)|^2.
\end{aligned}
$$

By adding and subtracting terms and the inequality $2ab \leqslant a^2/\varepsilon^2 + \varepsilon^2 b^2$, we have that

$$
\begin{aligned}
&|X_{u_1}(t) - X_{u_2}(t)|^2 + \mathbb{E}_t[\int_t^T |Z_{u_1}(s) - Z_{u_2}(s)|^2 ds] + \mathbb{E}_t[\int_t^T \|\ell_{u_1,t} - \ell_{u_2,s}\|_\nu^2 ds] \\
\leqslant& \frac{1}{\varepsilon^2} C \int_t^T \mathbb{E}_t\Big[\int_I \int_{\mathbb{R}} |f(s,x,X_{u_1}(s),Z_{u_1}(s),\ell_{u_1,s}(\cdot)) - f(s,x,X_{u_2}(s),Z_{u_2}(s),\ell_{u_2,s}(\cdot))|^2 \mu_{y,s}(dx) dy\Big] ds \\
&+ \frac{1}{\varepsilon^2} C \int_t^T \mathbb{E}_t\Big[\int_I \Big|\int_{\mathbb{R}} f(s,x,X_{u_1}(s),Z_{u_1}(s),\ell_{u_1,s}(\cdot))\mu_{y,s}(dx)\Big|^2 |G(u_1,y) - G(u_2,y)|^2 dy\Big] ds \\
&+ \mathbb{E}_t|X_{u_1}(T) - X_{u_2}(T)|^2 + \varepsilon^2 C \mathbb{E}_t[\int_t^T |X_{u_1}(s) - X_{u_2}(s)|^2 ds].
\end{aligned}
$$

Using the Lipschitz property of driver f and the solution $\Phi = (X, Z, \ell(\cdot)) \in \mathcal{M}$, we have

$$
\begin{aligned}
&\int_t^T \mathbb{E}_t\Big[\int_I \int_{\mathbb{R}} |f(s,x,X_{u_1}(s),Z_{u_1}(s),\ell_{u_1,s}(\cdot)) - f(s,x,X_{u_2}(s),Z_{u_2}(s),\ell_{u_2,s}(\cdot))|^2 \mu_{y,s}(dx) dy\Big] ds \\
\leqslant& \mathbb{E}_t[\int_t^T |X_{u_1}(s) - X_{u_2}(s)|^2 ds] + \mathbb{E}_t[\int_t^T |Z_{u_1}(s) - Z_{u_2}(s)|^2 ds] + \mathbb{E}_t[\int_t^T \|\ell_{u_1,t} - \ell_{u_2,s}\|_\nu^2 ds],
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_t^T \mathbb{E}\Big[\int_I \Big|\int_{\mathbb{R}} f(s,x,X_{u_1}(s),Z_{u_1}(s),\ell_{u_1,s}(\cdot))\mu_{y,s}(dx)\Big|^2 |G(u_1,y) - G(u_2,y)|^2 dy\Big] ds \\
\leqslant& C \int_t^T (1 + \sup_{y \in I} \mathbb{E}X_y^2(t) + \mathbb{E}X_{u_1}^2(s) + \mathbb{E}\|\ell_{u_1,s}\|_\nu + \mathbb{E}Z_{u_1}^2(s)) \Big(\int_I |G(u_1,y) - G(u_2,y)|^2 dy\Big) ds \\
\leqslant& C \int_t^T \int_I |G(u_1,y) - G(u_2,y)|^2 dy ds.
\end{aligned}
$$

By taking appropriate ε, we can have for some constant C,

$$
\begin{aligned}
&\mathbb{E}[\sup_{t \in [0,T]} |X_{u_1}(t) - X_{u_2}(t)|^2] + \mathbb{E}[\int_0^T |Z_{u_1}(s) - Z_{u_2}(s)|^2 ds] + \mathbb{E}[\int_0^T \|\ell_{u_1,t} - \ell_{u_2,s}\|_\nu^2 ds] \\
&\leqslant C\mathbb{E}|X_{u_1}(T) - X_{u_2}(T)|^2 + C \int_0^T \int_I |G(u_1,y) - G(u_2,y)|^2 dy ds + C \int_0^T \mathbb{E}|X_{u_1}(s) - X_{u_2}(s)|^2 ds.
\end{aligned}
$$

By noticing that for any $s \in [0,T]$,

$$\begin{aligned}\mathbb{E}|X_{u_1}(s) - X_{u_2}(s)|^2 \leqslant &\mathbb{E}[\sup_{t\in[0,T]} |X_{u_1}(t) - X_{u_2}(t)|^2] + \mathbb{E}[\int_0^T |Z_{u_1}(s) - Z_{u_2}(s)|^2 ds] \\ &+ \mathbb{E}[\int_0^T \|\ell_{u_1,t} - \ell_{u_2,s}\|_\nu^2 ds].\end{aligned}$$

It then follows by Gronwall inequality that for some new constant C,

$$\begin{aligned}&\mathbb{E}[\sup_{t\in[0,T]} |X_{u_1}(t) - X_{u_2}(t)|^2] + \mathbb{E}[\int_0^T |Z_{u_1}(s) - Z_{u_2}(s)|^2 ds] + \mathbb{E}[\int_0^T \|\ell_{u_1,t} - \ell_{u_2,s}\|_\nu^2 ds] \\ &\quad\leqslant C\mathbb{E}|X_{u_1}(T) - X_{u_2}(T)|^2 + C\int_I |G(u_1,y) - G(u_2,y)|^2 dy.\end{aligned}$$

Thus point (i) is proved. Taking the infimum over all random variables $X_{u_1}(T)$ and $X_{u_2}(T)$ such that $\mathcal{L}(X_{u_1}(T)) = \mathcal{L}(\xi_{u_1})$ and $\mathcal{L}(X_{u_2}(T)) = \mathcal{L}(\xi_{u_2})$, and combining this with the definition of $\mathcal{W}_{2,T}(\mu_{u_1}, \mu_{u_2})$, Assumption 5.4 and Assumption 5.5, we obtain continuity and Lipschitz continuity, respectively. ☐

Remark 5.9. *Note that even when the intensity measures $\nu_u, u \in I$ are different, the (Lipschitz) continuity results in Proposition 5.8 remain true. The continuity assumption on terminal conditions guarantees the continuity of the third component of solution. Under the canonical coupling, by proceeding as in the above proof, we get*

$$\begin{aligned}&\mathbb{E}[\sup_{t\in[0,T]} |X_{u_1}(t) - X_{u_2}(t)|^2] + \mathbb{E}[\int_0^T |Z_{u_1}(s) - Z_{u_2}(s)|^2 ds] + \mathbb{E}[\int_0^T \|\ell_{u_1,t}\circ\Phi_{u_1}^{-1} - \ell_{u_2,s}\circ\Phi_{u_2}^{-1}\|_\nu^2 ds \\ &\quad\leqslant C\mathbb{E}|X_{u_1}(T) - X_{u_2}(T)|^2 + C\int_I |G(u_1,y) - G(u_2,y)|^2 dy,\end{aligned}$$

where the measure ν is the canonical measure defined in the canonical coupling. We can regard $\ell\circ\Phi^{-1}$ as the third component of the solution.

We now study the stability of our graphon mean-field BSDE. That is, for a sequence of graphons G_n converging to some limit graphon G, in the sense of cut norm $\|G_n - G\|_\square \to 0$, we prove that the corresponding solution of the graphon mean-field BSDE converges in some sense (specified in the following theorem), and the law of the X component also converges in an integral sense of the Wasserstein distance $W_{2,T}$ on I.

Theorem 5.10. *Let (X, Z, ℓ) and (X^n, Z^n, ℓ^n) be the solutions of (5.1) associated with graphons G and G_n, terminal condition ξ and ξ^n, respectively. Suppose that f satisfies Assumption 5.2. Then we have*

$$\begin{aligned}&\mathbb{E}\Big[\int_I \big(\sup_{t\in[0,T]} |X_u^n(t) - X_u(t)|^2 + \int_0^T |Z_u^n(t) - Z_u(t)|^2 dt + \int_0^T \|\ell_{u,t}^n - \ell_{u,t}\|_{\nu_u}^2 dt\big) du\Big] \\ &\quad\leqslant C\Big[\int_I \mathbb{E}|\xi_u - \xi_u^n|^2 du + \|G - G_n\|_\square\Big].\end{aligned} \tag{5.10}$$

If $\|G_n - G\|_\Box \to 0$ *and* $\mathbb{E}[\int_I |\xi_u - \xi_u^n|^2 du] \to 0$ *as* $n \to \infty$*, it follows that*

$$\mathbb{E}\Big[\int_I \Big(\sup_{t\in[0,T]} |X_u^n(t) - X_u(t)|^2 + \int_0^T |Z_u^n(t) - Z_u(t)|^2 dt + \int_0^T \|\ell_{u,t}^n - \ell_{u,t}\|_{\nu_u}^2 dt\Big) du\Big] \to 0, \tag{5.11}$$

and consequently

$$\int_I \mathcal{W}_{2,T}(\mathcal{L}(X_u), \mathcal{L}(X_u^n)) \to 0. \tag{5.12}$$

Proof. Let $\mu_{u,s}^n$ be the law of $X_u^n(s)$. Similarly as in the proof of Proposition 5.8, for any $u \in I$, applying Itô's formula to $|X_u(t) - X_u^n(t)|$ and taking conditional expectation, we get

$$\begin{aligned}
&\mathbb{E}_t|X_u(t) - X_u^n(t)|^2 + \mathbb{E}_t\Big[\int_t^T |Z_u(s) - Z_u^n(s)|^2 ds\Big] + \mathbb{E}_t\Big[\int_t^T \|\ell_{u,t} - \ell_{u,s}^n\|_\nu^2 ds \\
\leqslant &2\mathbb{E}_t\Big[\int_t^T (X_u(s) - X_u^n(s)) \int_I \Big(\int_{\mathbb{R}} G(u,y) f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx) \\
&\quad - \int_{\mathbb{R}} G_n(u,y) f(s,x,X_u^n(s),Z_u^n(s),\ell_{u,s}^n(\cdot))\mu_{y,s}^n(dx)\Big) dy ds\Big] + \mathbb{E}_t|X_u(T) - X_u^n(T)|^2.
\end{aligned}$$

By adding and subtracting terms, we obtain

$$\begin{aligned}
&|X_u(t) - X_u^n(t)|^2 + \mathbb{E}_t\Big[\int_t^T |Z_u(s) - Z_u^n(s)|^2 ds\Big] + \mathbb{E}_t\Big[\int_t^T \|\ell_{u,t} - \ell_{u,s}^n\|_\nu^2 ds \\
\leqslant &\frac{1}{\varepsilon^2} C \int_t^T \mathbb{E}_t\Big[\int_I \Big|\int_{\mathbb{R}} f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx)\Big|^2 |G(u,y) - G_n(u,y)|^2 dy\Big] ds \\
+ &\frac{1}{\varepsilon^2} C \int_t^T \mathbb{E}_t\Big[\int_I \int_{\mathbb{R}} \big|f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot)) - f(s,x,X_u^n(s),Z_u^n(s),\ell_{u,s}^n(\cdot))\big|^2 G_n^2(u,y)\mu_{y,s}(dx)dy\Big] ds \\
+ &\frac{1}{\varepsilon^2} C \int_t^T \mathbb{E}_t\Big[\int_I \Big|\int_{\mathbb{R}} f(s,x,X_u^n(s),Z_u^n(s),\ell_{u,s}^n(\cdot)) G_n(u,y)[\mu_{y,s} - \mu_{y,s}^n](dx)\Big|^2 dy\Big] ds \\
+ &\mathbb{E}_t|X_u(T) - X_u(T)|^2 + \varepsilon^2 C \mathbb{E}_t\Big[\int_t^T |X_u(s) - X_u(s)|^2 ds\Big].
\end{aligned} \tag{5.13}$$

Denote the first three terms on the right hand side by $\mathcal{I}_{u,1}^n(t)$, $\mathcal{I}_{u,2}^n(t)$ and $\mathcal{I}_{u,3}^n(t)$ respectively. By using the Lipschitz property of f, the property of $(X, Z, \ell) \in \mathcal{M}$, we have for all $u \in I$,

$$\begin{aligned}
&C \int_0^T \mathbb{E}\Big[\Big|\int_{\mathbb{R}} f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx)\Big|^2\Big] \\
&\leqslant 2C \int_0^T \mathbb{E}\big[1 + |X_u(s)|^2 + |X_y(s)|^2 + |Z_u(s)|^2 + \|\ell_{u,s}\|_{\nu_u}^2\big] \\
&\leqslant 2CT + C \sup_{u\in I} \mathbb{E}\Big[\sup_{t\in[0,T]} |X_u(t)|^2 + \int_0^T |Z_u(s)|^2 ds + \int_0^T \|\ell_{u,s}\|_{\nu_u}^2 ds\Big] = O(1).
\end{aligned}$$

From the above inequality, we infer for all $u \in I$,

$$\mathcal{I}^n_{u,1}(0) \leqslant \frac{1}{\varepsilon^2} C \int_I \big(G(u,y) - G_n(u,y)\big)^2 dy.$$

By the equivalence between the cut norm and the L^1 operator norm of a graphon, we have

$$\int_I \mathcal{I}^n_{u,1}(0) du \leqslant \frac{1}{\varepsilon^2} C \int_I \int_I \big(G(u,y) - G_n(u,y)\big)^2 dy du \leqslant \frac{1}{\varepsilon^2} C \|G - G_n\|_{\square}.$$

Then again by the Lipschitz property of f and Remark 5.1, we have

$$\int_I \mathcal{I}^n_{u,2}(t) du \leqslant \frac{1}{\varepsilon^2} C \int_t^T \int_I \big(\mathbb{E}_t\big[|X_u(s) - X^n_u(s)|^2 + |Z_u(s) - Z^n_u(s)|^2 + \|\ell_{u,s} - \ell^n_{u,s}\|^2_{\nu_u}\big]\big) du ds,$$

and

$$\int_I \mathcal{I}^n_{u,3}(t) du \leqslant \frac{1}{\varepsilon^2} C \int_t^T \int_I (\mathcal{W}_{2,s}(\mu_y, \mu^n_y))^2 dy.$$

Similarly, by taking an appropriate ε and applying Gronwall inequality, we get with some new constant C

$$\begin{aligned}\int_I \mathbb{E}[\sup_{t\in[0,T]} |X_u(t) - X^n_u(t)|^2] du \leqslant & C\Big[\int_I \mathbb{E}|X_u(T) - X^n_u(T)|^2 du \\ & + \|G - G_n\|_{\square} + \int_0^T \int_I (\mathcal{W}_{2,s}(\mu_y, \mu^n_y))^2 dy ds\Big].\end{aligned}$$

Now by inserting the above inequality to (5.13) and finding another appropriate ε, we can obtain

$$\begin{aligned}&\mathbb{E}\Big[\int_I \big(\sup_{t\in[0,T]} |X^n_u(t) - X_u(t)|^2 + \int_0^T |Z^n_u(t) - Z_u(t)|^2 dt + \int_0^T \|\ell^n_{u,t} - \ell_{u,t}\|^2_{\nu_u} dt\big) du\Big] \\ &\leqslant C\Big[\int_I \mathbb{E}|X_u(T) - X^n_u(T)|^2 du + \|G - G_n\|_{\square} + \int_0^T \int_I (\mathcal{W}_{2,s}(\mu_y, \mu^n_y))^2 dy ds\Big].\end{aligned} \tag{5.14}$$

Notice that by the definition of $\mathcal{W}_{2,s}$, it is clear that

$$\begin{aligned}&\int_I (\mathcal{W}_{2,s}(\mu_y, \mu^n_y))^2 dy \leqslant \\ &\quad \mathbb{E}\Big[\int_I \big(\sup_{t\in[0,T]} |X^n_u(t) - X_u(t)|^2 + \int_0^T |Z^n_u(t) - Z_u(t)|^2 dt + \int_0^T \|\ell^n_{u,t} - \ell_{u,t}\|^2_{\nu_u} dt\big) du\Big].\end{aligned}$$

Hence again by Gronwall inequality and our assumptions, we can conclude that

$$\begin{aligned}&\mathbb{E}\Big[\int_I \big(\sup_{t\in[0,T]} |X^n_u(t) - X_u(t)|^2 + \int_0^T |Z^n_u(t) - Z_u(t)|^2 dt + \int_0^T \|\ell^n_{u,t} - \ell_{u,t}\|^2_{\nu_u} dt\big) du\Big] \\ &\leqslant C\Big[\int_I \mathbb{E}|X_u(T) - X^n_u(T)|^2 du + \|G - G_n\|_{\square}\Big].\end{aligned}$$

Thus Inequality (5.10) is proved. Results (5.11) and (5.12) follow from the convergence assumptions $\|G_n - G\|_\square \to 0$ and $\mathbb{E}[\int_I |\xi_u - \xi_u^n|^2 du] \to 0$ as $n \to \infty$. For the last result, since the laws of ξ and ξ^n are given, we can use the coupling arguments as in the proof of Proposition 5.8. Taking the infimum over ξ_u and ξ_u^n with $\mathcal{L}(\xi_u) = \Lambda_u$ and $\mathcal{L}(\xi_u^n) = \Lambda_u^n$, for each $u \in I$, it follows that

$$\int_I \mathcal{W}_{2,T}(\mathcal{L}(X_u), \mathcal{L}(X_u^n)) \leqslant C\Big[\int_I (\mathcal{W}_2(\Lambda_u, \Lambda_u^n))^2 du + \|G - G_n\|_\square\Big].$$

The proof is complete. $\square$

We next provide an example where the convergence of graphons, i.e., $\|G_n - G\|_\square \to 0$ as $n \to \infty$, is well known and Theorem 5.10 can be applied.

Example 5.11. *For a size n adjacency matrix A, we define the associated step graphon G_A as:*

$$G_A(u, v) := A_{ij}, \qquad \text{for} \quad (u, v) \in I_i^n \times I_j^n,$$

where $I_i^n := ((i-1)/n, i/n]$, for $i = 2, \dots, n$ and $I_1^n := [0, 1/n]$. Let ζ^n be the adjacency matrix of an Erdös-Rényi random graph $\mathcal{G}(n, p_n)$. If $p_n = p$ is fixed as $n \to \infty$, then it is well known that, as $n \to \infty$, the associated graphon G_{ζ^n} converges in cut norm to the constant graphon $G \equiv p$.

We now provide another stability result which provides the convergence of graphon mean-field BSDEs in the space $\mathcal{M}$.

Proposition 5.12. *With the same notation and under the same assumptions as in Theorem 5.10, we have*

$$\begin{aligned}
&\sup_{u \in I} \mathbb{E}\Big[\sup_{t \in [0,T]} |X_u^n(t) - X_u(t)|^2 + \int_0^T |Z_u^n(t) - Z_u(t)|^2 dt + \int_0^T \|\ell_{u,t}^n - \ell_{u,t}\|_{\nu_u}^2 dt\Big] \\
&\quad \leqslant C\Big[\sup_{u \in I} \mathbb{E}|\xi_u - \xi_u^n|^2 + \|G - G_n\|_{\infty \to \infty}\Big] \to 0.
\end{aligned}$$

Consequently,

$$\mathcal{W}_{2,T}^{\mathcal{M}}(\mathcal{L}(X), \mathcal{L}(X^n)) \to 0.$$

Furthermore, given the law of terminal conditions $\mathcal{L}(\xi_u) = \Lambda_u$ and $\mathcal{L}(\xi_u^n) = \Lambda_u^n$ for each $u \in I$, we have explicitly

$$\mathcal{W}_{2,T}^{\mathcal{M}}(\mathcal{L}(X), \mathcal{L}(X^n)) \leqslant C\big[(\mathcal{W}_2^{\mathcal{M}}(\Lambda, \Lambda^n))^2 + \|G - G_n\|_{\infty \to \infty}\big].$$

Proof. The proof is similar as the proof of Theorem 5.10. We highlight only the difference. By definition, we have

$$\mathcal{I}_{u,1}^n(0) \leqslant \frac{1}{\varepsilon^2} C \int_I \big(G(u, y) - G_n(u, y)\big) dy \leqslant \frac{1}{\varepsilon^2} C \|G - G_n\|_{\infty \to \infty}.$$

We then take the supremum of each term of the right hand side of (5.14), so that for all $u \in I$,

$$\begin{aligned}&\mathbb{E}\Big[\sup_{t\in[0,T]} |X_u^n(t) - X_u(t)|^2 + \int_0^T |Z_u^n(t) - Z_u(t)|^2 dt + \int_0^T \|\ell_{u,t}^n - \ell_{u,t}\|_{\nu_u}^2 dt\Big] \\ &\leqslant C\Big[\sup_{u\in I} \mathbb{E}|X_u(T) - X_u^n(T)|^2 + \|G - G_n\|_{\infty\to\infty} + \int_0^T \sup_{u\in I}(\mathcal{W}_{2,s}(\mu_y, \mu_y^n))^2 ds\Big].\end{aligned}$$

Thus we can put the supremum on the left hand side, that is

$$\begin{aligned}&\sup_{u\in I}\mathbb{E}\Big[\sup_{t\in[0,T]} |X_u^n(t) - X_u(t)|^2 + \int_0^T |Z_u^n(t) - Z_u(t)|^2 dt + \int_0^T \|\ell_{u,t}^n - \ell_{u,t}\|_{\nu_u}^2 dt\Big] \\ &\leqslant C\Big[\sup_{u\in I} \mathbb{E}|X_u(T) - X_u^n(T)|^2 + \|G - G_n\|_{\infty\to\infty} + \int_0^T \sup_{u\in I}(\mathcal{W}_{2,s}(\mu_y, \mu_y^n))^2 ds\Big].\end{aligned}$$

Then everything follows similarly as in the proof of Theorem 5.10. □

5.3 Convergence of interacting particle systems to graphon mean-field BSDEs

Consider a sequence of N particle graphon interacting systems with $N \in \mathbb{N}$. We prove that under proper assumptions, the sequence of particle systems converges to the graphon BSDE system, and the convergence rate is also precised.

We define the corresponding N-coupled BSDE system (for $i = 1, \dots, N$):

$$\begin{aligned}X_i^N(t) &= \xi_i^N + \int_t^T \frac{1}{N}\sum_{j=1}^N \zeta_{ij}^N f(s, X_j^N(s), X_i^N(s), Z_i^N(s), \ell_s^{N,i}(\cdot))ds - \int_t^T Z_i^N(s) d\widehat{W}_i(s) \\ &\quad - \int_t^T \int_E \ell_s^{N,i}(e)\widetilde{\widehat{N}}_i(ds, de), \quad t \in [0,T] \\ X_i^N(T) &= \xi_i^N,\end{aligned} \tag{5.15}$$

where $\widehat{W}_i := W_{\frac{i}{N}}$ are i.i.d. Brownian motions, and $\widehat{N}_i(dt, de) = N_{\frac{i}{N}}(dt, de)$ are independent Poisson random measures. We assume that $\xi_i^N \in L^2(\mathcal{F}_T)$ for all $i = 1, \dots, N$. Hereby, $\zeta_{ij}^N : \mathbb{N} \times \mathbb{N} \mapsto \mathbb{R}_0^+$ is symmetric, describing the strength of interaction between particle i and j. The graphon G can be regarded as the limit of ζ_{ij}^N as $N \to \infty$.

Similarly as before, we define the space

$$\mathcal{M}^N := \{\Phi^N := \{(X_i, Z_i, \ell_i(\cdot)) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{H}^2_{\nu_i^N}\}_{i=1}^N, \text{ such that}$$

$$\|\Phi^N\|_{\mathcal{M}^N} := \max_{i=1,\dots,N}\big(\mathbb{E}[\sup_{t\in[0,T]} |X_i(t)|^2] + \mathbb{E}[\int_0^T |Z_i(t)|^2 dt + \mathbb{E}[\int_0^T \|\ell_{i,t}\|_{\nu_i}^2 dt]\big)^{1/2} < \infty\},$$

where $\nu_i^N := \nu_{i/N}$.

We first provide a uniqueness theorem for the N-coupled BSDE system (5.15).

Theorem 5.13. *Let f satisfy Assumption 5.2. Then the N-coupled BSDE system (5.15) admits a unique solution $\Phi^N \in \mathcal{M}^N$.*

Proof. The proof is similar to that of the corresponding mean-field BSDE. We first establish the contraction property and the convergence of Picard iterative sequence. Here for notation convenience, we drop the superscript N for the system. We set $\Phi^0 = \{\Phi_i^0\}_{i=1}^N$ with $\Phi_i^0 = (0,0,0)$ for all $i = 1, \dots, N$, and define the iterative map $\Phi^{n+1} := \Psi(\Phi^n)$ at step $n \in \mathbb{N}$, where $\Phi^n = \{(X_i^n, Z_i^n, \ell_i^n)\}_{i=1}^N$ is defined by the solution of the following iterative equation:

$$\begin{aligned} X_i^n(t) &= \xi_i + \int_t^T \frac{1}{N}\sum_{j=1}^N \zeta_{ij} f(s, X_j^{n-1}(s), X_i^n(s), Z_i^n(s), \ell_{i,s}^n(\cdot))ds - \int_t^T Z_i^n(s)dW_i(s) \\ &\quad - \int_t^T \int_E \ell_{i,s}^n(e)\widetilde{N}_i(ds,de), \quad t \in [0,T], \\ X_i^n(T) &= \xi_i. \end{aligned} \tag{5.16}$$

Note that at each iteration step $n \geqslant 1$ and for each $i \leqslant N$, the existence and uniqueness of $(X_i^n, Z_i^n, \ell_i^n) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{H}^2_{\nu_i^N}$ is established by classical results since the driver is Lipschitz, see [183].

It is sufficient to prove that (5.15) admits a unique solution in $\mathcal{M}^N$. For convenience, denote $\|\cdot\|_i := \|\cdot\|_{\nu_i^N}$. We show that Ψ is a contraction in $\mathcal{M}^N$. As before, let $\overline{X}_i^n(t) := X_i^n(t) - X_i^{n-1}(t)$, $\overline{Z}_i^n(t) := Z_i^n(t) - Z_i^{n-1}(t)$ and $\overline{\ell}_t^{i,n} := \ell_t^{i,n} - \ell_t^{i,n-1}$. For $r > 0$, applying Itô's formula to $e^{rs}|\overline{X}_i^n(s)|^2$ between 0 and T, $n \geqslant 1$ and taking expectation, it follows that,

$$\begin{aligned} &\mathbb{E}[r\int_0^T e^{rs}(\overline{X}_i^n(s))^2 ds + \int_0^T e^{rs}(\overline{Z}_i^n(s))^2 ds + \int_0^T e^{rs}\|\overline{\ell}_s^n\|_i^2 ds] \\ \leqslant &\mathbb{E}\big[2\int_0^T e^{rs}\overline{X}_i^n(s)\frac{1}{N}\sum_{j=1}^N \zeta_{ij}\big(f(s, X_j^{n-1}(s), X_i^n(s), Z_i^n(s), \ell_{i,s}^n(\cdot)) \\ &\qquad - f(s, X_j^{n-2}(s), X_i^{n-1}(s), Z_i^{n-1}(s), \ell_s^{n-1,i}(\cdot))\big)ds\big]. \end{aligned}$$

Similarly let

$$\widetilde{A} := \frac{1}{N}\sum_{j=1}^N \zeta_{ij}(f(s, X_j^{n-1}(s), \Phi_i^n(s)) - f(s, X_j^{n-2}(s), \Phi_i^{n-1}(s)).$$

The Lipschitzness of the driver f and boundedness of ζ_{ij} imply that for some constant C_0,

$$\widetilde{A}^2 \leqslant \frac{1}{N}\sum_{j=1}^N C_0(|\overline{X}_j^n(s)|^2 + |\overline{X}_i^n(s)|^2 + |\overline{Z}_i^n(s)|^2 + \|\overline{\ell}_s^n\|_i^2).$$

Then by Cauchy-Schwarz inequality, we have for any $\varepsilon > 0$,

$$\begin{aligned}
&\mathbb{E}[r\int_0^T e^{rs}(\overline{X}_i^n(s))^2 ds + \int_0^T e^{rs}(\overline{Z}_i^n(s))^2 ds + \int_0^T e^{rs}\|\overline{\ell}_s^n\|_i^2 ds] \\
\leqslant & C_0^2\varepsilon^2\mathbb{E}[\int_0^T e^{rs}(\frac{1}{N}\sum_{j=1}^N(|\overline{X}_j^n(s)|^2 + |\overline{X}_i^n(s)|^2 + |\overline{Z}_i^n(s)|^2 + \|\overline{\ell}_s^n\|_i^2))ds] \\
&+ \frac{1}{\varepsilon^2}\mathbb{E}[\int_0^T e^{rs}|\overline{X}_i^n(s)|^2 ds] \\
\leqslant & C_0^2\varepsilon^2\{\mathbb{E}[\int_0^T e^{rs}(|\overline{X}_i^n(s)|^2 + |\overline{Z}_i^n(s)|^2 + \|\overline{\ell}_s^{i,n}\|_u^2)ds + \sup_i \mathbb{E}[\int_0^T |\overline{X}_i^{n-1}(s)|^2 ds]\} \\
&+ \frac{1}{\varepsilon^2}\mathbb{E}[\int_0^T e^{rs}|\overline{X}_i^n(s)|^2 ds].
\end{aligned}$$

By choosing r and ε such that $r - C_0^2\varepsilon^2 - \frac{1}{\varepsilon^2} > C_0^2\varepsilon^2$ and $1 - C_0^2\varepsilon^2 > 0$ at the same time, we obtain the contraction inequality with certain constant $\alpha < 1$:

$$\max_{i=1,\dots,N}\mathbb{E}[\int_0^T e^{rs}(|\overline{X}^n|^2 + |\overline{Z}^n|^2 + \|\overline{\ell}^{i,n}\|_i^2)ds] \leqslant \alpha\sup_{u\in I}\mathbb{E}[\int_0^T e^{rs}(|\overline{X}^{n-1}|^2 + |\overline{Z}^{n-1}|^2 + \|\overline{\ell}^{i,n-1}\|_i^2)ds].$$

We get further the contraction inequality in $\mathcal{M}^N$,

$$\max_{i=1,\dots,N}\mathbb{E}[\int_0^T (|\overline{X}^n|^2 + |\overline{Z}^n|^2 + \|\overline{\ell}^{i,n}\|_i^2)ds] \leqslant \alpha\sup_{u\in I}\mathbb{E}[\int_0^T (|\overline{X}^{n-1}|^2 + |\overline{Z}^{n-1}|^2 + \|\overline{\ell}^{i,n-1}\|_i^2)ds].$$

Therefore the map Ψ is a contraction in $\mathcal{M}^N$ and it has a unique fixed point, which is denoted by Φ^N. Now by taking a limit in the iterating equation (5.3), we conclude that Φ^N is the unique solution of (5.15). $\square$

With a mild regularity assumption on the terminal value and the interaction terms ζ_{ij}^N, we have the following convergence result.

Assumption 5.6. *For a given graphon G, we say that $\zeta^N := \{\zeta_{ij}^N\}_{i,j\in[N]}$ satisfies the regularity assumption with graphon G if either:*

(i) $\zeta_{ij}^N = G(\frac{i}{N}, \frac{j}{N})$*;*

(ii) $\zeta_{ij}^N = \text{Bernoulli}\big(G(\frac{i}{N}, \frac{j}{N})\big)$ *independently for all $1 \leqslant i \leqslant j \leqslant N$ and independent of $\{W_u, N_u, \xi_u : u \in I\}$.*

For notation simplicity, we let all ν_u be a common measure ν. But notice that all following results hold for different ν_u.

Theorem 5.14. *Let Assumptions 5.2 and 5.5 be fullfilled. Suppose that ζ^N satisfies the regularity assumption 5.6 with graphon G, G is Lipschitz continuous and the terminal conditions ξ^N and ξ satisfy*

$$\max_{i=1,\dots,N}\mathbb{E}|\xi_i^N-\xi_{\frac{i}{N}}|^2=O(N^{-1}).$$

Then the unique solution Φ^N of (5.15) converges to the unique solution of (5.1) with the convergence rate $1/\sqrt{N}$ and

$$\begin{aligned}&\max_{i=1,\dots,N}\mathbb{E}\left[\sup_{t\in[0,T]}|X_i^N(t)-X_{\frac{i}{N}}(t)|^2+\int_0^T|Z_i^N(t)-Z_{\frac{i}{N}}(t)|^2dt+\int_0^T\|\ell_t^{i,N}-\ell_t^{\frac{i}{N}}\|_\nu^2dt\right]\\&\quad\leqslant CN^{-1}+C\max_{i=1,\dots,N}\mathbb{E}|\xi_i^N-\xi_{\frac{i}{N}}|^2=O(N^{-1}),\end{aligned}\tag{5.17}$$

for all $N\in\mathbb{N}$ and some constant C. Furthermore, for $\kappa_t^N=\frac{1}{N}\sum_{i=1}^N\delta_{X_i^N(t)}$ and $\kappa_t=\int_I\mathcal{L}(X_u(t))du$, we have

$$\sup_{t\in[0,T]}\mathbb{E}\big[(\mathcal{W}_2(\kappa_t^N,\kappa_t))^2\big]\leqslant CN^{-1/2}.\tag{5.18}$$

Proof. For convenience, we denote by $\overline{X}_i(t):=X_i^N(t)-X_{\frac{i}{N}}(t)$, $\overline{Z}_i(t):=Z_i^N(t)-Z_{\frac{i}{N}}(t)$, $\overline{\ell}_s^i(\cdot):=\ell_t^{i,N}(\cdot)-\ell_t^{\frac{i}{N}}(\cdot)$ and $\overline{\xi}_i:=\xi_i^N-\xi_{\frac{i}{N}}$. By applying Itô's formula to $|\overline{X}_s|^2$ on $[t,T]$, we have

$$\begin{aligned}&\overline{X}_i(t)^2+\int_t^T\overline{Z}_i(s)^2ds+\int_t^T\|\overline{\ell}_s^i\|_\nu^2ds\\&=\overline{X}_i(T)^2+2\int_t^T\overline{X}_i(s)(\frac{1}{N}\sum_{j=1}^N\zeta_{ij}^Nf(s,X_j^N(s),\Phi_i^N(s))\\&\quad-\int_I\int_{\mathbb{R}}G(\frac{i}{N},y)f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{y,s}(dx)dy)ds-2\int_t^T\overline{X}_i(s)\overline{Z}_i(s)dW_s\\&\quad-2\int_t^T\int_E\overline{X}_i(s-)\overline{\ell}_s^i(e)\widetilde{N}(ds,de)\\&\leqslant a^2\int_t^T(\frac{1}{N}\sum_{j=1}^N\zeta_{ij}^Nf(s,X_j^N(s),\Phi_i^N(s))-\int_I\int_{\mathbb{R}}G(\frac{i}{N},y)f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{y,s}(dx)dy)^2ds\\&\quad\frac{1}{a^2}\int_t^T\overline{X}_i(s)^2ds+|\overline{\xi}_i|^2-2\int_t^T\overline{X}_s\overline{Z}_sdW_s-2\int_t^T\int_E\overline{X}_{s-}\overline{\ell}_s(e)\widetilde{N}(ds,de).\end{aligned}\tag{5.19}$$

We now put attention on the driver difference term. We will analyze the integrand for fixed

$s \in [t,T]$. Taking conditional expectation given $\mathcal{F}_t$, and adding and subtracting terms, we get

$$\begin{aligned}
&\mathbb{E}_t\Big|\frac{1}{N}\sum_{j=1}^N \zeta_{ij}^N f(s, X_j^N(s), \Phi_i^N(s)) - \int_I\int_{\mathbb{R}} G(\frac{i}{N}, y) f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{y,s}(dx)dy\Big|^2 \\
&\leqslant 3\mathbb{E}_t\Big|\frac{1}{N}\sum_{j=1}^N \zeta_{ij}^N (f(s, X_j^N(s), \Phi_i^N(s)) - f(s, X_{\frac{j}{n}}(s), \Phi_{\frac{i}{n}}(s)))\Big|^2 \\
&\quad + 3\mathbb{E}_t\Big|\frac{1}{N}\sum_{j=1}^N (\zeta_{ij}^N f(s, X_{\frac{j}{n}}(s), \Phi_{\frac{i}{n}}(s)) - \int_{\mathbb{R}} G(\frac{i}{N}, \frac{j}{n}) f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{\frac{j}{n},s}(dx))\Big|^2 \\
&\quad + 3\mathbb{E}_t\Big|\frac{1}{N}\sum_{j=1}^N \int_{\mathbb{R}} G(\frac{i}{N}, \frac{j}{n}) f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{\frac{j}{n},s}(dx) - \int_I\int_{\mathbb{R}} G(\frac{i}{N}, y) f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{y,s}(dx)dy\Big|^2 \\
&=: 3(\mathcal{I}_1^N(s) + \mathcal{I}_2^N(s) + \mathcal{I}_3^N(s)).
\end{aligned}$$

For the first term $\mathcal{I}_1^N(s)$, by Lipschitz continuity of f and boundedness of ζ_{ij}^N, we have

$$\mathcal{I}_1^N(s) \leqslant C_0(\max_{i=1,\dots,N} \mathbb{E}_t[|\overline{X}_i(s)|^2] + \mathbb{E}_t[|\overline{X}_i(s)|^2 + |\overline{Z}_i(s)|^2 + \|\overline{\ell}_s^i\|_\nu^2]). \tag{5.20}$$

For the second term $\mathcal{I}_2^N(s)$, we first analyse $\mathbb{E}[\mathcal{I}_2^N(s)]$. By the independence of $\{\zeta_{ij}^N\}$ and $\{X_{\frac{i}{n}}\}$, we have

$$\begin{aligned}
\mathbb{E}[\mathcal{I}_2^N(s)] = \frac{1}{N^2}\sum_{\mathcal{A}_{i,j,k}} &\mathbb{E}[(\zeta_{ij}^N f(s, X_{\frac{j}{n}}(s), \Phi_{\frac{i}{n}}(s)) - \int_{\mathbb{R}} G(\frac{i}{N}, \frac{j}{n}) f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{\frac{j}{n},s}(dx)) \\
&(\zeta_{ik}^N f(s, X_{\frac{k}{n}}(s), \Phi_{\frac{i}{n}}(s)) - \int_{\mathbb{R}} G(\frac{i}{N}, \frac{j}{n}) f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{\frac{k}{n},s}(dx))] \\
\leqslant\ & \frac{1}{N^2}|\mathcal{A}_{i,j,k}| C_0 \sup_{u\in I}\mathbb{E}[X_u(s)^2] \leqslant \frac{C_0}{N},
\end{aligned} \tag{5.21}$$

where $\mathcal{A}_{i,j,k}$ is the set of triplets $\{i,j,k\}$ in $[n]^3$ such that $j = k \neq i$, and its cardinality is of order $O(N)$. The last inequality follows from the uniform boundedness of the second moment of X_u on $[0,T]$.

We now estimate the expectation of the third term $\mathbb{E}[\mathcal{I}_3^N(s)]$. By adding and subtracting terms, we get

$$\begin{aligned}
\mathbb{E}[\mathcal{I}_3^N(s)] &\leqslant 2\mathbb{E}\Big|\int_I\int_{\mathbb{R}} [G(\frac{i}{N}, \frac{[Ny]}{N}) - G(\frac{i}{n}, y)] f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{\frac{[Ny]}{N},s}(dx)dy\Big|^2 \\
&\quad + 2\mathbb{E}\Big|\int_I\int_{\mathbb{R}} G(\frac{i}{N}, y) f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{\frac{[Ny]}{N},s}(dx)dy - \int_I\int_{\mathbb{R}} G(\frac{i}{N}, y) f(s,x,\Phi_{\frac{i}{n}}(s))\mu_{y,s}(dx)dy\Big|^2 \\
&\quad \leqslant \frac{C_0}{N^2},
\end{aligned} \tag{5.22}$$

where the last inequality comes from the Lipschitz property of f, the uniform boundedness of the second moments of $X_u(s)$ for all $u \in I$, the property of $\Phi \in \mathcal{M}$, Remark 5.1 and Proposition 5.8 (iii).

Coming back to (5.19) and taking conditional expectation given $\mathcal{F}_t$, denoted by $\mathbb{E}_t$, we have

$$\begin{aligned}
&|\overline{X}_i(t)|^2 + \int_t^T \mathbb{E}_t[\overline{Z}_i(s)^2]ds + \int_t^T \mathbb{E}_t[\|\overline{\ell}_s^i\|_\nu^2]ds \\
&\leqslant \mathbb{E}_t[|\overline{\xi}_i|^2] + 3a^2C_0 \int_t^T \mathbb{E}_t[\max_{i=1,\dots,N} |\overline{X}_i(s)|^2]ds + 3a^2C_0 \int_t^T \mathbb{E}_t[(|\overline{Z}_i(s)|^2 + \|\overline{\ell}_s^i\|_\nu^2)]ds \\
&\quad + (3a^2C_0 + \frac{1}{a^2}) \int_t^T \mathbb{E}_t[|\overline{X}_i(s)|^2]ds + \int_t^T (\mathcal{I}_2^N(s) + \mathcal{I}_3^N(s))ds.
\end{aligned} \tag{5.23}$$

By choosing $a^2 < 1/(3C_0)$, we obtain

$$\max_{i=1,\dots,N} |\overline{X}_i(t)|^2 \leqslant (6a^2C_0 + \frac{1}{a^2})\mathbb{E}_t[\int_t^T \max_{i=1,\dots,N} |\overline{X}_i(s)|^2 ds] + \int_t^T (\mathcal{I}_2^N(s) + \mathcal{I}_3^N(s))ds + \max_{i=1,\dots,N} \mathbb{E}_t[|\overline{\xi}_i|^2]. \tag{5.24}$$

We then apply Gronwall's inequality to get that

$$\max_{i=1,\dots,N} |\overline{X}_i(t)|^2 \leqslant C \int_t^T (\mathcal{I}_2^N(s) + \mathcal{I}_3^N(s))ds + C \max_{i=1,\dots,N} \mathbb{E}_t[|\overline{\xi}_i|^2],$$

for all $t \in [0, T]$. So we have further

$$\begin{aligned}
\max_{i=1,\dots,N} \mathbb{E}[\sup_{t\in[0,T]} |\overline{X}_i(t)|^2] \leqslant& C \int_0^T \mathbb{E}[\mathcal{I}_2^N(s) + \mathcal{I}_3^N(s)]ds + C \max_{i=1,\dots,N} \mathbb{E}_t[|\overline{\xi}_i|^2] \\
\leqslant& \frac{C}{N} + C \max_{i=1,\dots,N} \mathbb{E}_t[|\overline{\xi}_i|^2].
\end{aligned}$$

Inserting the above equation into (5.23) and choosing again $a^2 < 1/(3C_0)$, we have in turn

$$\mathbb{E}_t[\int_t^T \overline{Z}_i(s)^2 ds] + \mathbb{E}_t[\int_t^T \|\overline{\ell}_s^i\|_\nu^2 ds] \leqslant \frac{C}{N} + C \max_{i=1,\dots,N} \mathbb{E}_t[|\overline{\xi}_i|^2].$$

Combining the above two formulas we obtain (5.17).

We now show (5.25). Denote $\bar{\kappa}_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_{\frac{i}{N}}(t)}$. By (5.17), it is easily seen than

$$\sup_{t\in[0,T]} \mathbb{E}[(\mathcal{W}_2(\kappa_t^N, \bar{\kappa}_t^N))^2] \leqslant \frac{C}{N}.$$

Then notice that the jump diffusion of each X_u is influenced only by the laws of others, which can be regarded as independent. Using the uniform second moment bound of X_u for all $u \in I$ on $[0, T]$, we can apply [53, Lemma A.1], and get

$$\sup_{t\in[0,T]} \mathbb{E}[(\mathcal{W}_2(\bar{\kappa}_t^N, \mathbb{E}\bar{\kappa}_t^N))^2] \leqslant \frac{C}{\sqrt{N}}.$$

Finally, since the graphon G is Lipschitz continuous, by Proposition 5.8, the Lipschitz continuity property of the law of X_u on u in the Wasserstein L^2 distance guarantees that

$$\sup_{t\in[0,T]} \mathbb{E}[(\mathcal{W}_2(\mathbb{E}\bar{\kappa}_t^N, \kappa_t))^2] \leqslant \frac{C}{N}.$$

Combining the above three estimations, we obtain (5.25). The proof is now complete. □

In a relaxed case where ζ^N is related to a sequence of graphons G^N instead of a fixed graphon G, we can also have similar convergence result and obtain the convergence rate.

Theorem 5.15. *Let Assumptions 5.2 and 5.5 be fullfilled. Suppose ζ^N satisfies the regularity Assumption 5.6 with graphon G^N. Then we have*

$$\max_{i=1,\dots,N} \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right]$$
$$\leqslant C\big(\max_{i=1,\dots,N} \mathbb{E}|\xi_i^N - \xi_{\frac{i}{N}}|^2 + \|G - G^N\|_{\infty\to\infty} + N^{-1}\big).$$

Proof. Let $(\widetilde{X}^N, \widetilde{Z}^N, \widetilde{\ell}^N)$ be the solution of (5.1) induced by the step graphon $\widetilde{G}^N$ obtained from G^N, i.e. $\widetilde{G}^N(u,v) := G(\frac{[Nu]}{N}, \frac{[Nv]}{N})$, and terminal value $\widetilde{\xi}_u^N := \xi_i^N$ for $u \in (\frac{i-1}{N}, \frac{i}{N}]$, $i = 1, 2, \dots, N$ and $\widetilde{\xi}_0^N = \xi_1^N$. Notice that for each $N \in \mathbb{N}$, the regularity assumption for ζ^N with G^N and with $\widetilde{G}^N$ are equivalent. By Theorem 5.14, we have

$$\max_{i=1,\dots,N} \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - \widetilde{X}_{\frac{i}{N}}^N(t)|^2 + \int_0^T |Z_i^N(t) - \widetilde{Z}_{\frac{i}{N}}^N(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \widetilde{\ell}_t^{\frac{i}{N},N}\|_\nu^2 dt\right]$$
$$\leqslant CN^{-1} + C\max_{i=1,\dots,N} \mathbb{E}|\xi_i^N - \widetilde{\xi}_{\frac{i}{N}}^N|^2 \leqslant CN^{-1},$$

Since now every $\widetilde{G}^N$ is a step graphon, we have a uniformly bounded constant C for the sequence $\{\widetilde{G}^N\}$. On the other hand, by Proposition 5.12, we have

$$\max_{i=1,\dots,N} \mathbb{E}\big[\sup_{t\in[0,T]} |\widetilde{X}_{\frac{i}{N}}^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |\widetilde{Z}_{\frac{i}{N}}^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\widetilde{\ell}_t^{\frac{i}{N},N} - \ell_{\frac{i}{N},t}\|_{\nu_{\frac{i}{N}}}^2 dt\big]$$
$$\leqslant C\big[\sup_{u\in I} \mathbb{E}|\widetilde{\xi}_u^N - \xi_u|^2 + \|G - G^N\|_{\infty\to\infty}\big].$$

Combining the above two formulas and the definition of $\widetilde{\xi}^N$, we obtain the desired result. □

The following is a direct corollary of Theorem 5.15 under some Lipschitz conditions.

Corollary 5.16. *Suppose ζ^N satisfies the regularity Assumption 5.6 with graphon G^N with $\|G - G_N\|_\square \leqslant \frac{C}{N}$, where $G^N(u,v) := G^N(\frac{[Nu]}{N}, \frac{[Nv]}{N})$ is a step graphon. Let Assumptions 5.2 be fullfilled.*

Suppose the graphon G is Lipschitz continuous, $u \mapsto \xi_u$ is Lipschitz continuous with respect to $L^2(\mathcal{F}_T)$, and the terminal condition ξ^N satisfies

$$\max_{i=1,\dots,N} \mathbb{E}|\xi_i^N - \xi_{\frac{i}{N}}|^2 = O(N^{-1}).$$

Then we have $\|G - G^N\|_{\infty\to\infty} \leqslant \frac{C}{N}$ and thus

$$\max_{i=1,\dots,N} \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right] \leqslant \frac{C}{N},$$

and

$$\sup_{t\in[0,T]} \mathbb{E}\big[(\mathcal{W}_2(\kappa_t^N, \kappa_t))^2\big] \leqslant CN^{-1/2}, \tag{5.25}$$

where $\kappa_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i^N(t)}$ and $\kappa_t = \int_I \mathcal{L}(X_u(t))du$.

We end this section by some remarks.

Remark 5.17. *In the case of a sequence of random graphons which converge in probability in cut norm, we can show that Theorem 5.15 can still be applied. In particular, consider the following example taken from [162]. Let $U_1, \dots, U_N$ be i.i.d. uniform random variables on $[0,1]$ and $U_{(1)}, \dots, U_{(N)}$ be their order statistics. For a given graphon G, if $i \neq j$, then connect vertices (i,j) with probability $G(U_{(i)}, U_{(j)})$. Denoting the underlying adjacency matrix by ζ^N, we have that the graphons G_{ζ^N} associated to ζ^N (defined as in Example 5.11) converge in probability in cut norm to G. Notice that by the boundedness of the graphon, $\mathbb{E}\|G_{\zeta^N} - G\|_\square \to 0$, and for each realisation of $U_1, \dots, U_N$, ζ^N satisfies the regularity assumption 5.6. Thus if the measures $\{\nu_u\}_{u\in I}$ are a common measure ν, the graphon G is continuous, $u \mapsto \xi_u$ is continuous with respect to $L^2(\mathcal{F}_T)$, and the terminal condition ξ^N satisfies $\max_{i=1,\dots,N} \mathbb{E}|\xi_i^N - \xi_{\frac{i}{N}}|^2 \to 0$. In this case, Theorem 5.15 still applies.*

Remark 5.18. *Note that with the following weaker terminal conditions on ξ^N and ξ,*

$$\frac{1}{N}\sum_{i=1}^N \mathbb{E}|\xi_i^N - \xi_{\frac{i}{N}}|^2 = O(N^{-1}),$$

we can obtain similar convergence results of the average type:

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^N \mathbb{E}&\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right] \\ &\leqslant CN^{-1} + C\frac{1}{N}\sum_{i=1}^N \mathbb{E}|\xi_i^N - \xi_{\frac{i}{N}}|^2 = O(N^{-1}),\end{aligned}$$

similar as in Theorem 5.14. Also similar as in Theorem 5.15, we can show that

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^N \mathbb{E}&\left[\sup_{t\in[0,T]} |X_i^N(t) - X_{\frac{i}{N}}(t)|^2 + \int_0^T |Z_i^N(t) - Z_{\frac{i}{N}}(t)|^2 dt + \int_0^T \|\ell_t^{i,N} - \ell_t^{\frac{i}{N}}\|_\nu^2 dt\right] \\ &\leqslant C\Big(\frac{1}{N}\sum_{i=1}^N \int_{[\frac{i-1}{N}, \frac{i}{N}]} \mathbb{E}|\xi_i^N - \xi_u|^2 du + \|G - G^N\|_\square + N^{-1}\Big).\end{aligned}$$

for some constant C.

5.4 Graphon dynamic risk measures

5.4.1 Definition and properties

In this section, we introduce the graphon dynamic risk measures induced by the solution of a graphon mean-field BSDE system (5.1) and study its properties. Similar to [183], for $T > 0$ representing a given maturity and ξ a financial position at time T, we interpret $\rho_{u,t}(\xi, T) := -X_u(t, \xi, T)$, as the risk measure of ξ at time t and position $u \in I$.

Definition 5.19. *Let $T > 0$ be a time horizon, for a terminal condition $\xi \in \mathcal{M}L^2(\mathcal{F}_T)$, we define*

$$\rho_{u,t}(\xi, T) := -X_u(t, \xi, T),$$

for each $u \in I$, where $\{X_u(t, \xi, T)\}_{u \in I}$ is the solution of the graphon mean-field BSDE system (5.1). Then $\rho_t(\xi, T) := \{\rho_{u,t}(\xi, T)\}_{u \in I}$ is called the graphon associated dynamic risk measure.

We now provide some properties of the above dynamic risk measures, under Assumption 5.2. Let $\mathcal{T}_0$ be the set of all stopping times τ such that $\tau \in [0, T]$ almost surely.

(i) **Consistency**: Let $\tau \in \mathcal{T}_0$ be a stopping time. Then for each time t smaller than τ, the risk measure associated with terminal value ξ at maturity T coincides with the risk-measure associated with maturity τ and terminal value $-\rho_\tau(\xi, T) = X_\tau(\xi, T)$, that is for all $t \in [0, \tau]$, a.s. $\rho_t(\xi, T) = \rho_t(-\rho_\tau(\xi, T), \tau)$. This property is guaranteed by the uniqueness result of the graphon mean-field BSDE system (Theorem 5.4).

(ii) **Continuity**: Let $\{\tau^\alpha, \alpha \in \mathbb{R}\}$ be a family of stopping times in $\mathcal{T}_0$ converging a.s. to a stopping time $\tau^{\alpha_0} \in \mathcal{T}_0$ as $\alpha \to \alpha_0$. Let $\{\xi^\alpha, \alpha \in \mathbb{R}\}$ be a sequence of random families such that for each $\alpha \in \mathbb{R}$, ξ^α is $\mathcal{F}_{\tau^\alpha}$-measurable and $\mathbb{E}[\operatorname{ess\,sup}_\alpha (\xi^\alpha)^2] < \infty$. Suppose also that ξ^α converges a.s. to a $\mathcal{F}_{\tau^{\alpha_0}}$-measurable random variable ξ as α tends to α_0. Then for each stopping time $\hat{\tau} < \tau^\alpha, \alpha \in \mathbb{R}$, the random variables $\rho_{\hat{\tau}}(\xi^\alpha, \tau^\alpha) \to \rho_{\hat{\tau}}(\xi, \tau^{\alpha_0})$ a.s. and the processes $\rho(\xi^\alpha, \tau^\alpha) \to \rho(\xi, \tau^{\alpha_0})$ in $\mathcal{M}\mathbb{S}^2$ when $\alpha \to \alpha_0$.
This property follows from [183, Proposition A.6]. Indeed, for each $u \in I$, by making some modifications in the proof of [183, Proposition A.6], we can show that $\rho_u(\xi^\alpha, \tau^\alpha) \to \rho_u(\xi, \tau^{\alpha_0})$ in $\mathbb{S}^2$ when $\alpha \to \alpha_0$.

(iii) **Homogeneity:** If f is positively homogeneous with respect to (x', x, z, ℓ), i.e., for $a > 0$, $f(t, ax', ax, az, al) = af(t, x', x, z, \ell)$, then the risk measure ρ is positively homogeneous with respect to ξ, that is, for all $\lambda \geqslant 0$, $t \in [0, T]$ and $\xi \in \mathcal{M}L^2(\mathcal{F}_T)$, we have $\rho_t(\lambda\xi, T) = \lambda\rho_t(\xi, T)$.

(iv) **Translation invariance:** If f depends only on $(t, x' - x, z, \ell)$, that is $f(t, x', x, z, \ell) = h(t, x' - x, z, \ell)$ for some function h, then the risk measure satisfies the translation invariance property: for any $\xi \in \mathcal{M}L^2(\mathcal{F}_T), t_0 \in [0, T]$ and $\xi' \in \mathcal{M}L^2(\mathcal{F}_{t_0})$,

$$\rho_t(\xi + \xi', T) = \rho_t(\xi, T) - \xi' \quad \text{for all } t \in [0, T].$$

Proof. Suppose $\xi' \in \mathcal{M}L^2(\mathcal{F}_{t_0})$ and recall that $\mu_{y,s}$ is the law of $X_y(s)$. Let $\{X_u, Z_u, \ell_u\}_{u\in I}$ be the solution of (for each $u \in I$)

$$\begin{aligned} X_u(t) = \xi_u &+ \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx)dyds \\ &- \int_t^T Z_u(s)dW_u(s) - \int_t^T \int_E \ell_{u,s}(e)\widetilde{N}_u(ds,de), \quad t \in [t_0,T], \\ X_u(T) = \xi_u. \end{aligned} \tag{5.26}$$

Then by assumption, we have for each pair $(u,v) \in I^2$,

$$\int_{\mathbb{R}} G(u,y) f(s,x,X_u(s),Z_u(s),\ell_{u,s}(\cdot))\mu_{y,s}(dx) = \int_{\mathbb{R}} G(u,y) f(s,x,X_u(s)+\xi'_u,Z_u(s),\ell_{u,s}(\cdot))\mu'_{y,s}(dx),$$

where $\mu'_{y,s}$ is the law of $X_y(s)+\xi'_y$. Hence we have

$$\begin{aligned} X_u(t)+\xi'_u =& \int_t^T \int_I \int_{\mathbb{R}} G(u,y) f(s,x,X_u(s)+\xi'_u,Z_u(s),\ell_{u,s}(\cdot))\mu'_{y,s}(dx)dyds \\ &- \int_t^T Z_u(s)dW_u(s) - \int_t^T \int_E \ell_{u,s}(e)\widetilde{N}_u(ds,de) + \xi_u + \xi'_u, \quad t \in [t_0,T], \\ X_u(T)+\xi'_u =& \xi_u + \xi'_u, \end{aligned}$$

which implies that $\{X_u+\xi'_u, Z_u, \ell_u\}_u$ is the solution of the (5.26) with terminal value $\xi_u+\xi'_u$. Hence, for all $u \in I$, we have

$$\rho_{u,t}(\xi_u+\xi'_u,T) = \rho_{u,t}(\xi_u,T) - \xi'_u,$$

as desired. □

From now on, we assume that f satisfies Assumption 5.3 and is non-decreasing on x'.

(v) **Monotonicity:** The risk measure is non-increasing with respect to the terminal value ξ, i.e., for each $T>0$ and each $\xi^1,\xi^2 \in \mathcal{M}L^2(\mathcal{F}_T)$, if $\xi^1 \geqslant \xi^2$ a.s., then a.s. $\rho_t(\xi^1,T) \leqslant \rho_t(\xi^2,T), 0 \leqslant t \leqslant T$. This property follows directly by the comparison Theorem 5.6.

(vi) **Convexity:** If f is concave with respect to (x,z,l), then the dynamic risk measure is convex, that is for any $\lambda \in [0,1]$ and $\xi^1,\xi^2 \in \mathcal{M}L^2(\mathcal{F}_T)$, we have

$$\rho_t(\lambda\xi^1+(1-\lambda)\xi^2,T) \leqslant \lambda\rho_t(\xi^1,T)+(1-\lambda)\rho_t(\xi^2,T).$$

Proof. By our assumption, it is clear that for each $u \in I$, $F_u(t,\mathcal{L}(.),x,z,\ell)$ is concave on (x,z,ℓ). But it is also concave in $\mathcal{L}$. For two family of measures $\mu^1,\mu^2 \in \mathcal{P}(\mathcal{M}\mathbb{S}^2)$, we have

$$\begin{aligned} \lambda F_u(\cdot,\mu^1,\cdot)+(1-\lambda)F_u(\cdot,\mu^2,\cdot) &= \int_I \int_{\mathbb{R}} G(u,v) f(x,\cdot)(\lambda\mu_v^1+(1-\lambda)\mu_v^2)(dx)dv \\ &\leqslant F_u(\cdot,\lambda\mu^1+(1-\lambda)\mu^2,\cdot). \end{aligned}$$

Let $\{X_u^1, Z_u^1, \ell_u^1\}_u$ and $\{X_u^2, Z_u^2, \ell_u^2\}_u$ be the solutions of the graphon mean-field BSDE system (5.1) with terminal values ξ^1 and ξ^2, respectively. For a given $\lambda \in [0,1]$ and each $u \in I$, denote $\widehat{X}_u := \lambda X_u^1 + (1-\lambda)X_u^2$, $\widehat{Z}_u := \lambda Z_u^1 + (1-\lambda)Z_u^2$, $\widehat{\ell}_u := \lambda \ell_u^1 + (1-\lambda)\ell_u^2$. Then we have

$$
\begin{aligned}
&\lambda F_u(t, \mathcal{L}(X_t^1), X_u^1, Z_u^1, \ell_u^1) + (1-\lambda)F_u(t, \mathcal{L}(X_t^2), X_u^2, Z_u^2, \ell_u^2) \\
\leqslant &F_u(t, \lambda\mathcal{L}(X_t^1) + (1-\lambda)\mathcal{L}(X_t^2), \widehat{X}_u, \widehat{Z}_u, \widehat{\ell}_u).
\end{aligned}
$$

Let $\{\widetilde{X}_u, \widetilde{Z}_u, \widetilde{\ell}_u\}_u$ be the solution of (5.1) with terminal value $\lambda\xi^1 + (1-\lambda)\xi^2$. Thus it follows by comparison Theorem 5.6 that for all $u \in I$ and $t \in [0,T]$,

$$
\widetilde{X}_u \geqslant \lambda X_u^1 + (1-\lambda)X_u^2,
$$

which implies the desired result. □

(vii) **No Arbitrage:** Suppose now that the strict inequality holds in Assumption 5.3, so that we can apply the strict comparison Theorem 5.7. It follows easily that the dynamic risk measure satisfies that for each $T > 0$ and each $\xi^1, \xi^2 \in \mathcal{M}L^2(\mathcal{F}_T)$, if $\xi^1 \geqslant \xi^2$ a.s. and $\rho_t(\xi^1, T) = \rho_t(\xi^2, T)$ a.s. on an event $A \in \mathcal{F}_t$, then $\xi^1 = \xi^2$ a.s. on A.

Example 5.20. *We consider the following examples for the integrand function f:*

(i) Let $f(t, x', x, z, \ell) = x' - x$. It is non-decreasing in x' and concave in x. In addition, it is Lipschitz in both x' and x, positively homogeneous in (x', x) and satisfies Assumption 5.3. Thus the associated risk measure satisfies all the properties above.

(ii) Let $f(t, x', x, z, \ell) = x' + x$. It is not a function of $x' - x$, thus the risk measure does not satisfies the translation invariance property.

(iii) Let $f(t, x', x, z, \ell) = e^{-|x'-x|}$. The risk measure only satisfies properties (i), (ii) and (iv).

5.4.2 Dual representation

We provide a dual representation for the expectation of the global dynamic risk measure induced by the graphon mean-field BSDE, when the interaction f is concave on (x', x, z, ℓ).

We first introduce some notation. For convenience, let F_u denote the driver of u component in the graphon mean-field system (5.1), i.e.,

$$
F_u(\omega, t, \mathcal{L}(X_t), x, z, \ell(\cdot)) := \int_I \int_{\mathbb{R}} G(u,y) f(s, x', x, z, \ell(\cdot)) \mu_{y,s}(dx') dy.
$$

For each (ω, t) and each $u \in I$, we denote by $(F_u)^*$ the Fenchel-Legendre transform (see e.g., [112]), defined as

$$
\begin{aligned}
(F_u)^*(\omega, t, \mathcal{L}(Y), \beta_u, \alpha_u^1, \alpha_u^2) := &\sup_{(X,x,z,\ell) \in L^{2,I}(\mathcal{F}_t) \otimes \mathbb{R}^2 \otimes L^2_{\nu_u}} \{F_u(\omega, t, \mathcal{L}(X), x, z, \ell) \\
&- \langle X, Y \rangle_{L^{2,I}} - \beta_u x - \alpha_u^1 z - \langle \alpha_u^2, \ell_u \rangle_{\nu_u}\}.
\end{aligned} \tag{5.27}
$$

For given processes (β, γ), we define

$$H_{t,s}^{\beta,\gamma} := \exp\{\int_t^s (\beta_y + \gamma_y) dy\}.$$

For each $u \in I$ and given predictable process $\alpha_u = (\alpha_u^1, \alpha_u^2)$, we let $\mathcal{Q}_u^\alpha$ denote the probability absolutely continuous with respect to $\mathbb{P}$, which admits $\Gamma^{\alpha_u,T}$ as density, where Γ^{α_u} is the solution of

$$d\Gamma^{\alpha_u,t} = \Gamma^{\alpha_u,t-}\big(\alpha_{u,t}^1 dW_u(t) + \int_E \alpha_{u,t}^2(e) d\widetilde{N}_u(dt, de)\big), \quad \Gamma^{\alpha_u,0} = 1. \tag{5.28}$$

Let $\mathcal{D}_{u,T}$ be the set of predictable processes α_u such that

- $\int_0^T (\alpha_{u,s}^1)^2 ds + \int_0^T \|\alpha_{u,s}^2\|_{\nu_u}^2 ds$ is bounded;
- $\alpha_{u,t}^2(y) > -1$ $\nu_u(dy)$-a.s. for all $t \in [0, T]$.

Define $\mathcal{D}_T^I$ to be the set of all family of processes $\alpha := \{\alpha_u\}_{u \in I}$ such that for each $u \in I$, $\alpha_u \in \mathcal{D}_{u,T}$. By [183, Proposition 3.1, 3.2], we know that for any $\alpha \in \mathcal{D}_T^I$, $\Gamma^{\alpha_u,t} > 0$ a.s. on $[0, T]$ and $(\Gamma^{\alpha_u,t})_{t \in [0,T]} \in \mathbb{S}^2$, for each $u \in I$.

In the following, we denote by $\gamma_t := (\gamma_t^{u,v})_{u,v \in I}$ for all $t \in [0, T]$. Let $\mathcal{A}_T^I$ be the set of families of processes $(\gamma_t, \beta_t, \alpha_t)_{t \in [0,T]}$, where $(\beta_t, \alpha_t)_{t \in [0,T]}$ are predictable and $(\gamma_t)_{t \in [0,T]}$ is progressively measurable, such that

- $(\alpha_t)_{t \in [0,T]}$ belongs to $\mathcal{D}_T^I$;
- For each $(u, v) \in I$, $(\Gamma_{v,t}^{\overline{\alpha}} e^{\int_0^t \overline{\gamma}_y^{u,v} dy})_{t \in [0,T]}$ belong to $\mathbb{H}^2$;
- For each $v \in I$, $\{(F_v)^*\big(t, \big(\frac{\Gamma_t^{\alpha_{v_1}} H_{0,t}^{\beta_{v_1},\gamma^{v,v_1}} \gamma_t^{v,v_1}}{\mathbb{E}[\Gamma_t^{\alpha_v} \int_I H_{0,t}^{\beta_v,\gamma^{v_1,v}} dv_1]}\big)_{v_1 \in I}, \beta_{v,t}, \alpha_{v,t}^1, \alpha_{v,t}^2(\cdot)\big)\}_{t \in [0,T]}$ belongs to $\mathbb{H}^2$.

We first provide some technical lemmas which will be used for the main duality theorem .

Lemma 5.21. *Suppose that f satisfies Assumption 5.3 and is non-decreasing in x'. Then for each (s, ω) and each $u \in I$, the effective domain of $(F_u)^*$ such that*

$$(F_u)^*(s, \mathcal{L}(Y), \beta_u, \alpha_u^1, \alpha_u^2) < +\infty$$

is included in the closed set $U_u \subseteq \mathbb{R}^2 \times L_{\nu_u}^2$ of all elements $(\beta_u, \alpha_u^1, \alpha_u^2)$ satisfying the following:

(i) β_u and α_u^1 are bounded by some constant C_1.

(ii) $\alpha_u^2 > -1$ and $|\alpha_u^2(y)| \leqslant C_2$ $\nu_u(dy)$ a.s. for some constant C_2.

Proof. By our assumptions, for each $u \in I$, the function F_u satisfies the same assumptions as in the classical case studied in [183, Lemma 5.4], thus the results follow by similar arguments. ☐

For each $t \in [0,T]$, a given random variable X in $L^{2,I}(\mathcal{F}_t)$ and each $u \in I$, we denote by $(\mathfrak{G}_u)^*$ the Fenchel-Legendre transform of $F_u(\cdot, \mathcal{L}(X), \cdot)$, defined as

$$(\mathfrak{G}_u)^*_{\beta_u,\alpha_u}(t,X) := \sup_{(x,z,\ell)\in\mathbb{R}^2\times L^2_{\nu_u}} \{F_u(t,\mathcal{L}(X),x,z,\ell) - \beta_u x - \alpha_u^1 z - \langle \alpha_u^2, \ell\rangle\}.$$

Further, we denote by $(\mathfrak{G}_u)^{**}_{\beta_u,\alpha_u}(t,Y)$ the Fenchel-Legendre transform of $(\mathfrak{G}_u)^*_{\beta_u,\alpha_u}(t,X)$, that is

$$(\mathfrak{G}_u)^{**}_{\beta_u,\alpha_u}(t,Y) := \sup_{X\in L^{2,I}(\mathcal{F}_T)} \{(\mathfrak{G}_u)^*_{\beta_u,\alpha_u}(t,X) - \langle X, Y\rangle_{L^{2,I}}\}.$$

Lemma 5.22. *For each $u \in I$, each (ω,t) and any given $(Y,\beta_u,\alpha_u) \in L^{2,I}(\mathcal{F}_t)\otimes\mathbb{R}^2\otimes L^2_{\nu_u}$ belonging to the effective domain of $(F_u)^*$, we have*

$$(\mathfrak{G}_u)^{**}_{\beta_u,\alpha_u}(t,Y) = (F_u)^*(\omega,t,\mathcal{L}(Y),\beta_u,\alpha_u^1,\alpha_u^2).$$

Proof. Obviously, one has

$$(\mathfrak{G}_u)^{**}_{\beta_u,\alpha_u}(t,Y) \leqslant (F_u)^*(\omega,t,\mathcal{L}(Y),\beta_u,\alpha_u^1,\alpha_u^2).$$

Suppose that (X_1,x_1,z_1,ℓ_1) attains the supremum of

$$F_u(t,\mathcal{L}(X),x,z,\ell) - \beta_u x - \alpha_u^1 z - \langle \alpha_u^2,\ell\rangle.$$

Then we have

$$F_u(t,\mathcal{L}(X_1),x_1,z_1,\ell_1) - \beta_u x_1 - \alpha_u^1 z_1 - \langle\alpha_u^2,\ell_1\rangle \leqslant (\mathfrak{G}_u)^*_{\beta_u,\alpha_u}(t,X_1).$$

Thus it follows that

$$F_u(t,\mathcal{L}(X_1),x_1,z_1,\ell_1) - \langle X_1, Y\rangle_{L^{2,I}} - \beta_u x_1 - \alpha_u^1 z_1 - \langle\alpha_u^2,\ell_1\rangle_{\nu_u} \leqslant (\mathfrak{G}_u)^{**}_{\beta_u,\alpha_u}(t,Y),$$

which implies that

$$(\mathfrak{G}_u)^{**}_{\beta_u,\alpha_u}(t,Y) \geqslant (F_u)^*(\omega,t,\mathcal{L}(Y),\beta_u,\alpha_u^1,\alpha_u^2).$$

Hence we get the desired equality. □

Lemma 5.23. *Assume that f is non-decreasing in x'. Then for any given $(\beta,\alpha)\in\mathbb{R}^{\otimes I}\times\mathbb{R}^{\otimes I}\times(L^2_{\nu_u})^{\otimes I}$, for each $t\in[0,T]$ and each $u\in I$, the effective domain of $(\mathfrak{G}_u)^{**}_{\beta_u,\alpha_u}(t,Y_u)$, i.e.,*

$$\{Y_u \in L^{2,I}(\mathcal{F}_t) : (\mathfrak{G}_u)^{**}_{\beta_u,\alpha_u}(s,Y_u) < +\infty\},$$

satisfies that $0 \leqslant Y_u \leqslant C$, for some positive constant C, $d\mathbb{P}\otimes d\lambda$ a.s. where λ is the Lebesgue measure on $\mathbb{R}$.

Proof. First note that since f is non-decreasing, we have that for all $u \in I$, F_u is also non-decreasing and thus $(\mathfrak{G}_u)^*_{\beta_u,\alpha_u}(s, X_s)$ is non-decreasing for each $u \in I$ and any $(\beta, \alpha) \in \mathbb{R}^{\otimes I} \times \mathbb{R}^{\otimes I} \times (L^2_{\nu_u})^{\otimes I}$. Thus we can drop the subscript u. For simplicity, we write $(\mathfrak{G}_u)^{**}_{\beta_u,\alpha_u}$ as $\mathfrak{G}^*$ and $(\mathfrak{G}_u)^*_{\beta_u,\alpha_u}$ as $\mathfrak{G}$.

We prove by contradiction. For each $s \in [0, T]$, suppose that $d\mathbb{P} \otimes d\lambda$ a.s. $0 \leqslant Y \leqslant C$, for some positive constant C, is not true. Denote by $A := \{\omega \times v : Y_v(\omega) < 0\}$ and $B := \{\omega \times v : Y_v(\omega) > C\}$. By the definition of $\mathfrak{G}^*$, for each $X \in L^{2,I}(\mathcal{F}_t)$ we have

$$\mathfrak{G}^*(s, Y) \geqslant \mathfrak{G}(s, X) - \langle X, Y\rangle = \mathfrak{G}(s, X) - \mathbb{E}[\int_I X_v Y_v dv].$$

We can take $X^n_v(\omega) := -nY_v \mathbb{1}_A(\omega)$. Clearly $X^n \leqslant 0$ $d\mathbb{P} \otimes d\lambda$ a.s., by the non-decreasing property of $\mathfrak{G}$ we have

$$\mathfrak{G}^*(s, Y) \geqslant \mathfrak{G}(s, 0) - \mathbb{E}[\int_I X^n_v Y_v dv] = \mathfrak{G}(s, 0) + n\int_A |Y_v(\omega)|^2 d(P \times \ell)(\omega, v).$$

Letting $n \to \infty$, we get $\mathfrak{G}^*(s, Y) = +\infty$, which gives the contradiction.

Notice that by the Lipschitz and non-decreasing property of $\mathfrak{G}$ and boundedness of the graphon, we have for some positive constant C,

$$\mathfrak{G}(s, X) \geqslant \mathfrak{G}(s, 0) + C\mathbb{E}[\int_I X_v dv].$$

Thus by taking $X^n_v(\omega) := -nY_v \mathbb{1}_B(\omega)$, similarly we have,

$$\begin{aligned}\mathfrak{G}^*(s, Y) \geqslant& \mathfrak{G}(s, 0) + C\mathbb{E}[\int_I X^n_v dv] - \mathbb{E}[\int_I X^n_v Y_v dv] \\ &= \mathfrak{G}(s, 0) + n\int_B Y_v(\omega)(Y_v(\omega) - C) d(\mathbb{P} \otimes \lambda)(\omega, v).\end{aligned}$$

Letting $n \to \infty$, we get the desired contradiction again. □

Example 5.24. *If f takes the form as in Example 5.20(i), that is $f(t, x', x, z, \ell) = x' - x$, then one can show that for each $u \in I$,*

$$(\mathfrak{G}_u)^*_{\beta_u}(t, X) = \int_I G(u, y)\mathbb{E}[X_y]dy,$$

and

$$(F_u)^*(t, \mathcal{L}(Y), \beta_u) = 0.$$

Indeed, by similar arguments as in the proof of Lemma 5.23, the effective domain is $(\mathbf{1}, \int_I G(u, y)\mathbb{E}[X_y]dy)$, where $\mathbf{1}$ is the set of all random variables Y in $L^{2,I}(\mathcal{F}_T)$ satisfying $Y = 1$ $d\mathbb{P} \otimes d\lambda$ a.s.. It is easily seen that F_u and $(F_u)^*$ satisfy the conjugacy relation (see e.g., [112]) and

$$F_u(t, \mathcal{L}(X), x) = (F_u)^{**}(t, \mathcal{L}(X), x).$$

Lemma 5.25. *Let $(\alpha^1_s, \alpha^2_s(\cdot))_{s\geqslant t}$ belong to $\mathcal{D}^I_T$, $(U^{u,v}(s))_{u,v\in I}$ and $(h_u)_{u\in I}$ be progressively measurable and bounded. Let $t \in [0,T]$. Then the following SDE system admits a solution $(V^{u,v})_{u,v\in I}$ uniformly in $\mathbb{S}^2$:*

$$\begin{aligned} dV^{u,v}(s) =& V^{u,v}(s)[\alpha^1_{v,s}dW_v(s) + \int_E \alpha^2_{v,s}(e)d\widetilde{N}_v(ds,de)] \\ &+ U^{u,v}(s)\mathbb{E}[\int_I V^{v,u}(s)h_{u,s}dv]ds, \quad t\leqslant s\leqslant T, \\ V^{u,v}(t) =& 1. \end{aligned}$$

Proof. This equation is a graphon mean field forward SDE system. Set $V_0(s)\equiv 1$ for $s\in[t,T]$ and define the following iteration sequence for $n\geqslant 1$,

$$V^{u,v}_n(s) = V^{u,v}_{n-1}(t) + \int_t^s V^{u,v}_{n-1}(y)dM_{v,y} + \int_t^s U^{u,v}(s)\mathbb{E}[\int_I V^{v,u}_{n-1}(y)h_{u,y}dv]dy,$$

where $dM_{v,y} = \alpha^1_{v,s}dW_v(s) + \int_E \alpha^2_{v,s}(r)d\widetilde{N}_v(ds,de)$.

Since $(\alpha^1_s, \alpha^2_s(\cdot))_{s\geqslant t}$ belong to $\mathcal{D}^I_T$, the quadratic variation in $[t,T]$ is bounded. By Doob L^2-inequality and Cauchy-Schwarz inequality, we have for some constant C,

$$\begin{aligned} \mathbb{E}[\sup_{t\leqslant s\leqslant T}|V^{u,v}_{n+1}(s) - V^{u,v}_n(s)|^2] \leqslant& C\mathbb{E}[\int_t^T |V^{u,v}_n(s) - V^{u,v}_{n-1}(s)|^2 ds] \\ &+ C\mathbb{E}[\int_t^T\int_I |V^{u,v}_n(s) - V^{u,v}_{n-1}(s)|^2 dvds]. \end{aligned}$$

Hence, we have further

$$D_{n+1} := \mathbb{E}[\sup_{\substack{(u,v)\in I^2\\ t\leqslant s\leqslant T}}|V^{u,v}_{n+1}(s) - V^{u,v}_n(s)|^2] \leqslant C\mathbb{E}[\int_t^T \sup_{\substack{(u,v)\in I^2\\ t\leqslant s\leqslant T}}|V^{u,v}_n(s) - V^{u,v}_{n-1}(s)|^2 ds].$$

We thus get, for any $n\geqslant 1$,

$$D_{n+1} \leqslant \frac{D_1 C^n T^n}{n!}.$$

Then consequently we have

$$\sum_{n=1}^{\infty}(D_n)^{1/2} < \infty,$$

which implies that uniformly in (u,v), $(V^{u,v}_n)_{n\geqslant 0}$ admits a uniform limit on $[t,T]$, which is a right-continuous process $V^{u,v}$. The family of processes $(V^{u,v})_{u,v\in I}$ solves the SDEs system and they are uniformly in $\mathbb{S}^2$. □

We provide now a dual representation theorem for the expectation of the integral of the graphon dynamic risk measure. We need the following lemma.

Lemma 5.26. *Let f be concave with respect to (x',x,z,ℓ) and non-decreasing in x'. Then given a family of processes $(X_s,\overline{\beta}_s,\overline{\alpha}_s)_{s\geqslant t}$ with X progressive, $\overline{\beta}$ bounded and $\overline{\alpha}\in\mathcal{D}_T^I$, there exists a family of progressively measurable processes $\overline{\gamma}_s=(\overline{\gamma}_s^{u,v})_{u,v\in I}$ such that for each $u,v\in I$, $(\Gamma_{v,s}^{\overline{\alpha}}e^{\int_t^s\overline{\gamma}_y^{u,v}dy})_{s\geqslant t}$ belongs to $\mathbb{H}^2$ and satisfies the following for each $v\in I$,*

$$
\begin{aligned}
(\mathfrak{G}_v)^*_{\overline{\beta}_{v,s},\overline{\alpha}_{v,s}}(s,X_s)-\int_I\frac{\mathbb{E}[\Gamma_s^{\overline{\alpha}^{v_1}}H_{t,s}^{\overline{\beta}_{v_1},\overline{\gamma}^{v,v_1}}\overline{\gamma}_s^{v,v_1}X_{v_1}(s)]}{\mathbb{E}[\Gamma_{v,s}^{\overline{\alpha}}\int_I H_{t,s}^{\overline{\beta}_v,\overline{\gamma}^{v_1,v}}dv_1]}dv_1&\\
=(\mathfrak{G}_v)^{**}_{\overline{\beta}_{v,s},\overline{\alpha}_{v,s}}\big(s,\big(\frac{\Gamma_s^{\overline{\alpha}^{v_1}}H_{t,s}^{\overline{\beta}_{v_1},\overline{\gamma}^{v,v_1}}\overline{\gamma}_s^{v,v_1}}{\mathbb{E}[\Gamma_{v,s}^{\overline{\alpha}}\int_I H_{t,s}^{\overline{\beta}_v,\overline{\gamma}^{v_1,v}}dv_1]}\big)_{v_1}\big),&
\end{aligned}
$$

where $(\cdot)_{v_1}$ means the family of random variables with $v_1\in I$, and Γ^{α_v} is defined by (5.28) with initial value $\Gamma^{\alpha_{v,t}}=1$.

Proof. For simplicity of notation, we drop the bar symbol over the processes. For any $(s,\omega)\in\Omega\times[0,T]$, f is Lipschitz, concave in (x',x,z,ℓ), we can deduce easily that for any $(s,\omega)\in\Omega\times[0,T]$ and any $u\in I$, F_v is Lipschitz and concave in $(\mathcal{L}(X),x,z,\ell)$. Indeed, F_v is Lipschitz and concave in (x,z,ℓ), and it suffices to prove the Lipschitzness and concavity in the first parameter. By Remark 5.1, we have

$$
\begin{aligned}
\big|F_v(\mathcal{L}(X_1),\cdot)-F_v(\mathcal{L}(X_2),\cdot)\big|&=\int_I\Big|\int_{\mathbb{R}}G(v,u)f(x,\cdot)(\nu_u^1-\nu_u^2)(dx)\Big|du\\
&\leqslant\int_I C(\mathbb{E}|X_1^u-X_2^u|^2)^{1/2}du\\
&\leqslant\|X_1-X_2\|_{L^{2,I}}.
\end{aligned}
$$

For any $\lambda\in[0,1]$, we have that

$$
\begin{aligned}
\lambda F_v(\mathcal{L}(X_1),\cdot)+(1-\lambda)F_v(\mathcal{L}(X_2),\cdot)&=\int_I\int_{\mathbb{R}}G(v,u)f(x,\cdot)(\lambda\nu_u^1+(1-\lambda)\nu_u^2)(dx)du\\
&=F_v(\lambda\mathcal{L}(X_1)+(1-\lambda)\mathcal{L}(X_2),\cdot).
\end{aligned}
$$

We can easily deduce that, for all $u\in I$, $(\mathfrak{G}_v)^*_{\overline{\beta}_{v,s},\overline{\alpha}_{v,s}}(s,X_s)$ is concave and Lipschitz continuous in $L^{2,I}(\mathcal{F}_t)$. Therefore it follows that for each s and $u\in I$, there exists $Y_{u,s}\in L^{2,I}(\mathcal{F}_t)$ such that

$$
(\mathfrak{G}_u)^*_{\overline{\beta}_{u,s},\overline{\alpha}_{u,s}}(s,X_s)-\mathbb{E}[\int_I X_v(s)Y_s^{u,v}dv]=(\mathfrak{G}_v)^{**}_{\overline{\beta}_{v,s},\overline{\alpha}_{v,s}}(s,Y_{u,s}).
$$

We next consider a dense countable subset $\widetilde{I}:=\{0=a_1<\cdots<a_i<a_{i+1}<\cdots\leqslant 1\}$ of I and let $Y_s^{u,v}=Y_{u,s}^{(i)}$ for all $v\in[a_i,a_{i+1})$, where we choose $Y_{u,s}^{(i)}=Y_s^{u,v^*}$ for certain $v^*\in[a_i,a_{i+1})$ such that $Y_s^{u,v^*}\in L^2(\mathcal{F}_s)$. Note that the maximizer $Y_{u,s}$ is not unique and we choose $Y_{u,s}=\sum_i\mathbb{1}_{[a_i,a_{i+1})}(v)Y_{u,s}^{(i)}$. Since $(L^2(\mathcal{F}_s))^{\otimes\widetilde{I}}\in L^{2,I}(\mathcal{F}_s)$ is a separable Hilbert space, by using the measurable selection theorem (see e.g., [199]), there exists a family of processes Y_u for each $u\in I$, such that $Y^{u,v}:[t,T]\mapsto L^2(\mathcal{F}_T)$ is measurable for each $v \in I$. Then by similar arguments as in [89, Lemma 3.4], we obtain the progressive measurability for $Y^{u,v}$ when $u \in I$ fixed.

Since f is non-decreasing, we have that for all $v \in I$, F_u is also non-decreasing and thus $(\mathfrak{G}_u)^*_{\overline{\beta}_{u,s},\overline{\alpha}_{u,s}}(s, X_s)$ is non-decreasing for each $u \in I$ and any $(\beta, \alpha) \in \mathbb{R}^{\otimes I} \times \mathbb{R}^{\otimes I} \times (L^2_{\nu_u})^{\otimes I}$. By Lemma 5.23, we have $Y_{u,s} \geqslant 0$ $d\mathbb{P} \otimes d\lambda$ a.s. and $Y_{u,s} \leqslant C$ $d\mathbb{P} \otimes d\lambda$ a.s. Let

$$U^{u,v}(s) = Y^{u,v}_s \exp\{-\int_t^s \beta_{v,y} dy\},$$

and,

$$h_{u,s} = \exp\{\int_t^s \beta_{u,y} dy\}.$$

Then by applying Itô's formula to $V^{u,v}(\Gamma^{\alpha_v})^{-1}$, we obtain

$$d(V^{u,v}(\Gamma^{\alpha_v})^{-1})_s = (\Gamma^{\alpha_v}_s)^{-1} e^{-\int_t^s \beta_{v,y} dy} Y^{u,v}_s \mathbb{E}[\int_I V^{v,u}(s) dv e^{\int_t^s \beta_{u,y} dy}] ds.$$

Let $\widetilde{V}^{u,v} := V^{u,v}(\Gamma^{\alpha_v})^{-1}$. Since changes in a null measured set in I do not change the $\|\cdot\|_{L^{2,I}}$ norm of $Y_{u,s}$, we can assume that for all $(u,v) \in I^2$, $0 \leqslant Y^{u,v}_s \leqslant C$ ($d\mathbb{P}$ a.s.). Therefore we have for all $(u,v) \in I^2$, $\widetilde{V}^{u,v} > 0$ a.s.. Thus we can choose for each $(s,\omega) \in [t,T] \times \Omega$, $\gamma^{u,v}_s(\omega) = \frac{d}{ds}(\log \widetilde{V}^{u,v}_s)(\omega)$, which is well-defined. Then $\gamma^{u,v}$ satisfies $e^{\int_t^s \gamma^{u,v}_y dy} = \widetilde{V}^{u,v}_s$. We obtain that

$$\gamma^{u,v}_s e^{\int_t^s \gamma^{u,v}_y dy} ds = (\Gamma^{\alpha_v}_s)^{-1} e^{-\int_t^s \beta_{v,y} dy} Y^{u,v}_s \mathbb{E}[\Gamma^{\alpha_{u,s}} \int_I e^{\int_t^s \gamma^{v,u}_y dy} e^{\int_t^s \beta_{u,y} dy} dv] ds.$$

Hence we have

$$\frac{\Gamma^{\alpha_v}_s H^{\beta_v,\gamma^{u,v}}_{t,s} \gamma^{u,v}_s}{\mathbb{E}[\Gamma^{\alpha_{u,s}} \int_I H^{\beta_u,\gamma^{v,u}}_{t,s} dv]} = Y^{u,v}_s, \qquad a.s,$$

and clearly $\Gamma^{\overline{\alpha}}_{v,s} e^{\int_t^s \overline{\gamma}^{u,v}_y dy} = V^{u,v}(s)$ belongs to $\mathbb{H}^2$. Note that the progressive measurability for $Y^{u,v}$ when $u \in I$ fixed, implies the progressive measurability of $\gamma^{u,v}$. The proof is now complete. ☐

We are now ready to provide the dual representation theorem.

Theorem 5.27 (Dual representation). *Suppose f satisfies Assumption 5.2 and 5.3. Moreover, suppose that f is concave with respect to (x', x, z, ℓ) and non-decreasing in x'. Then, for each $t \in [0,T]$, the expectation of the convex risk-measure ρ_t has the following representation : for each $\xi \in \mathcal{M}L^2(\mathcal{F}_T)$,*

$$\mathbb{E}[\int_I \rho_{v,t}(\xi, T) dv] = \sup_{(\gamma,\beta,\alpha) \in \mathcal{A}^I_T} \{\int_I \mathbb{E}^{\mathcal{Q}^\alpha_v}[-(\int_I H^{\beta_v,\gamma^{u,v}}_{t,T} du)\xi_v] dv - \int_I \zeta_{v,t}(\gamma, \beta, \alpha, T) dv\}, \tag{5.29}$$

where the function ζ, which is called the penalty function, is defined for each T and $(\gamma, \beta, \alpha) \in \mathcal{A}^I_T$ by

$$\zeta_{v,t}(\gamma,\beta,\alpha,T) = \int_t^T \mathbb{E}^{\mathcal{Q}^\alpha_v}\big[(\int_I H^{\beta_v,\gamma^{u,v}}_{t,s} du)(F_v)^*\big(s, (\frac{\Gamma^{\alpha_{v_1}}_s H^{\beta_{v_1},\gamma^{v,v_1}}_{t,s} \gamma^{v,v_1}_s}{\mathbb{E}[\Gamma^{\alpha_v}_s \int_I H^{\beta_v,\gamma^{v_1,v}}_{t,s} dv_1]})_{v_1}, \beta_{v,s}, \alpha^1_{v,s}, \alpha^2_{v,s}(\cdot)\big)\big] ds,$$

with Q_v^α the absolutely continuous probability measure with respect to $\mathbb{P}$ admitting density Γ^{α_v}, which is defined by (5.28) with initial value $\Gamma^{\alpha_v,t}=1$. Moreover, there exists $(\overline{\gamma},\overline{\beta},\overline{\alpha})\in\mathcal{A}_T^I$ attaining the supremum in (5.29). In particular, for each $v\in I$,

$$\mathbb{E}[\rho_{v,t}(\xi,T)]=\mathbb{E}^{\mathcal{Q}^{\overline{\alpha}_v}}[-(\int_I H_{t,T}^{\overline{\beta}_v,\overline{\gamma}^{u,v}}du)\xi_v]-\zeta_{v,t}(\overline{\gamma},\overline{\beta},\overline{\alpha},T).$$

Proof. For each family of predictable processes $(\gamma,\beta,\alpha)\in\mathcal{A}_T^I$, we apply Itô's formula to

$$\int_I\int_I H_{t,s}^{\beta_v,\gamma^{u,v}}X_v(s)dvdu$$

between t and T. We obtain

$$\begin{aligned}\int_I X_v(t)dv&=\int_I\int_I H_{t,T}^{\beta_v,\gamma^{u,v}}\xi_v dvdu-\int_I\int_I\int_t^T H_{t,s}^{\beta_v,\gamma^{u,v}}Z_v(s)dW_v(s)dvdu\\&\quad-\int_I\int_I\int_t^T\int_E H_{t,s}^{\beta_v,\gamma^{u,v}}\ell_{v,s}(e)\tilde{N}_v(ds,de)dvdu\\&\quad+\int_t^T\int_I\int_I H_{t,s}^{\beta_v,\gamma^{u,v}}\big[-\beta_{v,s}X_v(s)-\gamma_s^{u,v}X_v(s)+F_v(s,\mathcal{L}(X),X_v(s),Z_v(s),\ell_{v,s}(\cdot))\big]dvduds\\&=\int_t^T\int_I(\int_I H_{t,s}^{\beta_v,\gamma^{u,v}}du)\big[-\beta_{v,s}X_v(s)-\alpha_{v,s}^1Z_v(s)-\langle\alpha_{v,s}^2,\ell_{v,s}\rangle+F_v(s,\mathcal{L}(X),X_v(s),Z_v(s),\ell_{v,s}(\cdot))\big]dvds\\&\quad-\int_t^T\int_I\int_I H_{t,s}^{\beta_v,\gamma^{u,v}}\gamma_s^{u,v}X_v(s)dvduds+\int_I\int_I H_{t,T}^{\beta_v,\gamma^{u,v}}\xi_v dvdu-\int_t^T\int_I dM^{\mathcal{Q}_v^\alpha}(s)dv,\end{aligned}$$

where we recall that Q_v^α is the absolutely continuous probability measure with respect to $\mathbb{P}$ admitting density Γ^{α_v}, which is defined by (5.28) with initial value $\Gamma^{\alpha_v,t}=1$, and

$$dM^{\mathcal{Q}_v^\alpha}(s)=(\int_I H_{t,s}^{\beta_v,\gamma^{u,v}}du)Z_v(s)dW_v^{\mathcal{Q}_v^\alpha}(s)+\int_E\big(\int_I H_{t,s}^{\beta_v,\gamma^{u,v}}du\big)\ell_{v,s}(e)\tilde{N}_v^{\mathcal{Q}_v^\alpha}(ds,de),$$

with $dW_v^{\mathcal{Q}_v^\alpha}(s)=dW_v(s)-\alpha_{v,s}^1ds$ is a Brownian motion under $\mathcal{Q}_v^\alpha$, and $\tilde{N}_v^{\mathcal{Q}_v^\alpha}(ds,de)=\tilde{N}_v(ds,de)-\alpha_{v,s}^2(e)\nu_v(de)ds$ is the $\mathcal{Q}_v^\alpha$-compensated Poisson random measure $N_v(\cdot,\cdot)$.

Then notice that

$$\begin{aligned}&\int_t^T\int_I\int_I\mathbb{E}^{\mathcal{Q}_v^\alpha}[H_{t,s}^{\beta_v,\gamma^{u,v}}\gamma_s^{u,v}X_v(s)]dvduds\\&\quad=\int_t^T\int_I\mathbb{E}^{\mathcal{Q}_u^\alpha}\big[(\int_I H_{t,s}^{\beta_u,\gamma^{v,u}}dv)(\int_I\frac{\mathbb{E}[\Gamma_s^{\alpha_v}H_{t,s}^{\beta_v,\gamma^{u,v}}\gamma_s^{u,v}X_v(s)]}{\mathbb{E}[\Gamma_s^{\alpha_u}\int_I H_{t,s}^{\beta_u,\gamma^{v,u}}dv]}dv)\big]duds]\\&\quad=\int_t^T\int_I\mathbb{E}^{\mathcal{Q}_v^\alpha}\big[(\int_I H_{t,s}^{\beta_v,\gamma^{u,v}}du)(\int_I\frac{\mathbb{E}[\Gamma_s^{\alpha_{v_1}}H_{t,s}^{\beta_{v_1},\gamma^{v,v_1}}\gamma_s^{v,v_1}X_{v_1}(s)]}{\mathbb{E}[\Gamma_s^{\alpha_v}\int_I H_{t,s}^{\beta_v,\gamma^{v_1,v}}dv_1]}dv_1)\big]dvds].\end{aligned}$$

For each $v \in I$, we compute the expectation under Q_v^α. Then by taking expectation on both sides, we get

$$\begin{aligned}\mathbb{E}[\int_I X_v(t)dv] &= \int_I \mathbb{E}^{Q_v^\alpha}[(\int_I H_{t,T}^{\beta_v,\gamma^{u,v}} du)\xi_v]dv \\ &+ \int_t^T \int_I \mathbb{E}^{Q_v^\alpha}\Big[(\int_I H_{t,s}^{\beta_v,\gamma^{u,v}} du)[-\beta_{v,s}X_v(s) - \alpha_{v,s}^1 Z_v(s) \\ &- \langle \alpha_{v,s}^2, \ell_{v,s}\rangle - \int_I \frac{\mathbb{E}[\Gamma_s^{\alpha_{v_1}} H_{t,s}^{\beta_{v_1},\gamma^{v,v_1}} \gamma_s^{v,v_1} X_{v_1}(s)]}{\mathbb{E}[\Gamma_s^{\alpha_v}\int_I H_{t,s}^{\beta_v,\gamma^{v_1,v}} dv_1]} dv_1 + F_v(s,\mathcal{L}(X), X_v(s), Z_v(s), \ell_{v,s}(\cdot))]\Big]dvds.\end{aligned}$$

By the definition of Fenchel-Legendre transform and Lemma 5.26, we have that

$$\begin{aligned}(\mathfrak{G}_u)^*_{\beta_{u,s},\alpha_{u,s}}(s,X_s) &- \int_I \frac{\mathbb{E}[\Gamma_s^{\alpha_{v_1}} H_{t,s}^{\beta_{v_1},\gamma^{v,v_1}} \gamma_s^{v,v_1} X_{v_1}(s)]}{\mathbb{E}[\Gamma_s^{\alpha_v}\int_I H_{t,s}^{\beta_v,\gamma^{v_1,v}} dv_1]} dv_1 \\ &\leqslant (F_v)^*(s, (\frac{\Gamma_s^{\alpha_{v_1}} H_{t,s}^{\beta_{v_1},\gamma^{v,v_1}} \gamma_s^{v,v_1}}{\mathbb{E}[\Gamma_s^{\alpha_v}\int_I H_{t,s}^{\beta_v,\gamma^{v_1,v}} dv_1]})_{v_1}, \beta_{v,s}, \alpha_{v,s}^1, \alpha_{v,s}^2(\cdot)).\end{aligned}$$

Then we obtain

$$\begin{aligned}\mathbb{E}[\int_I X_v(t)dv] &\leqslant \inf_{(\gamma,\beta,\alpha)\in\mathcal{A}_T^I} \int_I \mathbb{E}^{Q_v^\alpha}[(\int_I H_{t,T}^{\beta_v,\gamma^{u,v}} du)\xi_v]dv \\ &+ \int_t^T \int_I \mathbb{E}^{Q_v^\alpha}\Big[(\int_I H_{t,s}^{\beta_v,\gamma^{u,v}} du)(F_v)^*\big(s, (\frac{\Gamma_s^{\alpha_{v_1}} H_{t,s}^{\beta_{v_1},\gamma^{v,v_1}} \gamma_s^{v,v_1}}{\mathbb{E}[\Gamma_s^{\alpha_v}\int_I H_{t,s}^{\beta_v,\gamma^{v_1,v}} dv_1]})_{v_1}, \beta_{v,s}, \alpha_{v,s}^1, \alpha_{v,s}^2(\cdot)\big)\Big]dvds.\end{aligned} \tag{5.30}$$

Let U_v be the set defined in Lemma 5.21. For any given law $\mathcal{L}(X_s)$, since F_v is concave in (x,z,ℓ) for all $v \in I$, we have the following conjugacy relation for all $v \in I$,

$$\begin{aligned}F_v(s,\mathcal{L}(X_s),x,z,\ell) &= \inf_{(\beta_v,\alpha^v)\in U_v}\{(\mathfrak{G}_v)^*_{\beta_v,\alpha^v}(s,X_s) + \beta_v x + \alpha_v^1 z + \langle \alpha_v^2, l\rangle\} \\ &= (\mathfrak{G}_v)^*_{\overline{\beta}_v,\overline{\alpha}_v}(s,X_s) + \overline{\beta}_v x + \overline{\alpha}_v^1 z + \langle \overline{\alpha}_v^2, l\rangle.\end{aligned} \tag{5.31}$$

The set U_v is strongly closed and convex in $\mathbb{R}^2 \times L^2_{\nu_v}$. Moreover U_v is bounded, thus it is compact. By some similar arguments as those in the proof of step 1 of [183, Lemma 5.5], we can conclude that there exists some $(\overline{\beta}_v, \overline{\alpha}_v) \in U_v$ satisfying (5.31).

Let $\mathcal{U}_v$ be the set of all triplets $(\gamma^v, \beta_v, \alpha_v)$ such that (β_v, α_v) are predictable and take values in U_v, and γ^v is progressively measurable and satisfies that, for all $u \in I$, $\big(\Gamma^{\alpha_u} e^{\int_t^s \gamma_y^{v,u} dy}\big)_{s\geqslant t}$ belongs to $\mathbb{H}^2$. Since $\mathbb{R}^2 \times L^2_{\nu_v}$ is Polish, we can apply the measurable selection theorem as in [183, Lemma 5.5], to assert that there exist predictable processes $(\overline{\beta}_{v,s}, \overline{\alpha}_{v,s}^1, \overline{\alpha}_{v,s}^2)_{s\geqslant t} \in \mathcal{U}_v$ such that a.s.

$$F_v(s,\mathcal{L}(X_s),X_v(s),Z_v(s),\ell_{v,s}) = (\mathfrak{G}_v)^*_{\overline{\beta}_{v,s},\overline{\alpha}_{v,s}}(s,X_s) + \overline{\beta}_{v,s}X_v(s) + \overline{\alpha}_{v,s}^1 Z_v(s) + \langle \overline{\alpha}_v^2, \ell_{v,s}\rangle. \tag{5.32}$$

Clearly for any $(\beta_s, \alpha_s)_{s\geqslant t}$ and $v \in I$, $(\mathfrak{G}_v)^*_{\beta_{v,s},\alpha_{v,s}}$ is also Lipschitz and concave on the space $L^{2,I}(\Omega)$. Therefore by Lemma 5.26 we have that for the processes $(X_s, \overline{\beta}_s, \overline{\alpha}_s)_{s\geqslant t}$ and all $v \in I$, there exists a family of progressively measurable processes $\overline{\gamma}_s = (\overline{\gamma}^{u,v}_s)_{u,v\in I}$ satisfying

$$(\mathfrak{G}_v)^*_{\overline{\beta}_{v,s},\overline{\alpha}_{v,s}}(s, X_s) - \int_I \frac{\mathbb{E}[\Gamma^{\overline{\alpha}^{v_1}}_s H^{\overline{\beta}_{v_1},\overline{\gamma}^{v,v_1}}_{t,s} \overline{\gamma}^{v,v_1}_s X_{v_1}(s)]}{\mathbb{E}[\Gamma^{\overline{\alpha}}_{v,s} \int_I H^{\overline{\beta}_v,\overline{\gamma}^{v_1,v}}_{t,s} dv_1]} dv_1$$
$$= (\mathfrak{G}_v)^{**}_{\overline{\beta}_{v,s},\overline{\alpha}_{v,s}}\big(s, \big(\frac{\Gamma^{\overline{\alpha}^{v_1}}_s H^{\overline{\beta}_{v_1},\overline{\gamma}^{v,v_1}}_{t,s} \overline{\gamma}^{v,v_1}_s}{\mathbb{E}[\Gamma^{\overline{\alpha}}_{v,s} \int_I H^{\overline{\beta}_v,\overline{\gamma}^{v_1,v}}_{t,s} dv_1]}\big)_{v_1}\big).$$

Notice that by Lemma 5.22

$$(\mathfrak{G}_v)^{**}_{\overline{\beta}_{v,s},\overline{\alpha}_{v,s}}\big(s, \big(\frac{\Gamma^{\overline{\alpha}^{v_1}}_s H^{\overline{\beta}_{v_1},\overline{\gamma}^{v,v_1}}_{t,s} \overline{\gamma}^{v,v_1}_s}{\mathbb{E}[\Gamma^{\overline{\alpha}}_{v,s} \int_I H^{\overline{\beta}_v,\overline{\gamma}^{v_1,v}}_{t,s} dv_1]}\big)_{v_1}\big)$$
$$= (F_v)^*\big(s, \big(\frac{\Gamma^{\overline{\alpha}^{v_1}}_s H^{\overline{\beta}_{v_1},\overline{\gamma}^{v,v_1}}_{t,s} \overline{\gamma}^{v,v_1}_s}{\mathbb{E}[\Gamma^{\overline{\alpha}}_{v,s} \int_I H^{\overline{\beta}_v,\overline{\gamma}^{v_1,v}}_{t,s} dv_1]}\big)_{v_1}, \overline{\beta}_{v,s}, \overline{\alpha}^1_{v,s}, \overline{\alpha}^2_{v,s}(\cdot)\big).$$

Let $\overline{\mathcal{U}}$ be the set $\overline{\mathcal{U}} := \mathcal{U}^{\otimes I}_v$. Together with (5.30), we finally obtain that

$$\begin{aligned}\mathbb{E}\Big[\int_I \rho_{v,t}(\xi, t) dv\Big] = \sup_{(\gamma,\beta,\alpha)\in\overline{\mathcal{U}}} \Big\{ &\int_I \mathbb{E}^{\mathcal{Q}^\alpha_v}\Big[-\Big(\int_I H^{\beta_v,\gamma^{u,v}}_{t,T} du\Big)\xi_v\Big] dv \\ - \int_t^T \int_I \mathbb{E}^{\mathcal{Q}^\alpha_v}\Big[\Big(\int_I H^{\beta_v,\gamma^{u,v}}_{t,s} du\Big)(F_v)^*\big(s, &\big(\frac{\Gamma^{\alpha_{v_1}}_s H^{\beta_{v_1},\gamma^{v,v_1}}_{t,s} \gamma^{v,v_1}_s}{\mathbb{E}[\Gamma^{\alpha_v}_s \int_I H^{\beta_v,\gamma^{v_1,v}}_{t,s} dv_1]}\big)_{v_1}, \beta_{v,s}, \alpha^1_{v,s}, \alpha^2_{v,s}(\cdot)\big)\Big] dv ds\Big\}.\end{aligned} \tag{5.33}$$

Then notice that by (5.32), as a sum of processes in $\mathbb{H}^2_T$, the process $(\mathfrak{G}_v)^*_{\overline{\beta}_{v,s},\overline{\alpha}_{v,s}}(s, X_s)$ belongs to $\mathbb{H}^2_T$ for all $v \in I$. Thus $(\overline{\gamma}, \overline{\beta}, \overline{\alpha}) \in \mathcal{A}^I_T$, which implies that the equality in (5.33) holds with $\overline{\mathcal{U}}$ replaced by $\mathcal{A}^I_T$. The proof is now complete. □

5.5 Concluding remarks

We have extended the standard framework of mean-field BSDEs with jumps to the graphon mean-field framework in order to capture heterogeneous interactions. We have studied the existence, uniqueness, and measurability of the solutions. We have proven that an interacting mean-field particle system with heterogeneous interactions converges to the graphon mean-field BSDEs system in a certain sense. Furthermore, we have provided comparison theorems for the graphon mean-field BSDEs with jumps. Analogous to the standard case, we introduced graphon dynamic risk measures, which are induced by the solution of a graphon mean-field BSDE system, and have explored some of their properties. We have also proven a dual representation theorem for the graphon dynamic risk measure in the convex case.

Future work will include a generalization of the driver f. We could allow it to depend on the law of the entire solution processes $(X_u, Z_u, \ell_u)_{u\in I}$, not only on the law of X_u. More ambitiously, we may anticipate that the interacting driver could take different forms for different pairs of interactions. For example, we could consider a blockwise model where $I = \bigcup_{i=1}^{K} I_i$ for some finite K. For a pair $(u, v) \in I_i \times I_j$, the interaction between X_u and X_v could be specified by $f_{ij} = f_{ji}$, where for each pair (i, j), f_{ij} satisfies the assumptions of this paper. The formal formulation of this blockwise graphon mean-field BSDE system could then be given by:

$$
\begin{aligned}
X_u(t) = \xi_u + \sum_{j=1}^{K} \int_t^T \int_{I_j} \int_{\mathbb{R}} G(u,y) f_{ij}(s, x, X_u(s), Z_u(s), \ell_{u,s}(\cdot)) \mu_{y,s}(dx) dy ds \\
- \int_t^T Z_u(s) dW_u(s) - \int_t^T \int_E \ell_{u,s}(e) \widetilde{N}_u(ds, de), \quad \text{for} \quad t \in [0,T] \quad \text{and} \quad u \in I.
\end{aligned}
$$

More generally, the driver f can even differ from pair to pair. Another interesting aspect is to attempt to make the heterogeneous interaction depend on the underlying network structure. A key assumption in this paper is that we can capture interaction through the labels. However, in some cases, heterogeneous interactions between pairs may depend on other network parameters, e.g., distances between pairs, node centrality, etc. Moreover, the underlying graph can be considered in various spaces, not limited to $\mathbb{R}^d$. Another direction is to try to obtain more accurate approximation for the solutions of particle systems around those of the limit graphon systems, in particular, to investigate the CLT results in the graphon mean field setting, extending those in the classical mean field setting [72]. Moreover getting some concentration bounds and concentration measures for the backward particle systems with jumps as those in [51, 54] for forward systems is left for future works.

Chapter 6

Stochastic Graphon Mean Field Games with Jumps and Approximate Nash Equilibria

This chapter is based on paper [6] in the publication list of Section 1.5.

Abstract. We study continuous stochastic games with inhomogeneous mean field interactions on large networks and explore their graphon limits. We consider a model with a continuum of players, where each player's dynamics involve not only mean field interactions but also individual jumps induced by a Poisson random measure. We examine the case of controlled dynamics, with control terms present in the drift, diffusion, and jump components. We introduce the graphon game model based on a graphon controlled stochastic differential equation (SDE) system with jumps, which can be regarded as the limiting case of a finite game's dynamic system as the number of players goes to infinity. Under some general assumptions, we establish the existence and uniqueness of Markovian graphon equilibria. We then provide convergence results on the state trajectories and their laws, transitioning from finite game systems to graphon systems. We also study approximate equilibria for finite games on large networks, using the graphon equilibrium as a benchmark. The rates of convergence are analyzed under various underlying graphon models and regularity assumptions.

Keywords: Graphon mean field games, Jump measures, Heterogenous interactions, Controlled dynamics, Approximate Nash equilbria.

Contents

5.1 Introduction **184**

5.2 Graphon mean-field BSDEs with jumps **185**

5.2.1 Notation and definitions 185

5.2.2 Graphons 186

5.2.3 Existence and uniqueness results 188

5.2.4 Comparison theorems 195
5.2.5 Continuity and stability results 197
5.3 Convergence of interacting particle systems to graphon mean-field BSDEs 204
5.4 Graphon dynamic risk measures 212
5.4.1 Definition and properties 212
5.4.2 Dual representation 214
5.5 Concluding remarks 223

6.1 Introduction

The study of mean field systems with homogeneous interaction dates back to the work of Boltzmann, Vlasov, McKean, and others (see e.g., [33, 154, 172]). The theory of mean field games (MFG), introduced by Lasry and Lions in [163] and Huang, Caines, and Malhamé [141, 142], has attracted considerable attention and been extensively studied in recent decades; see, in particular, the recent book [83] and references therein. As both large n player and limiting models are quite tractable, the MFG theory has developed a diverse and broad range of applications. However, despite some MFG models incorporating heterogeneity in individual characteristics, the framework for MFG theory remains largely confined to games with homogeneous interactions, where all players are symmetrically exchangeable.

The study of stochastic games on large networks presents significant challenges, as various n player networks may yield different limits when n approaches infinity, particularly in the context of games on sparse networks (see, for example, [120, 161]). Analyzing games on large networks or those with heterogeneous interactions often relies on a tractable limiting (continuum) model, which can, in turn, offer insights into large finite games.

Recently, the use of graphons has emerged as a model to analyse heterogeneous interaction in mean field systems and heterogeneous game theory, see in particular [60, 78, 79]. Graphons have been developed by Lovász et al., see e.g. [67, 68, 170], as a natural continuum limit object for large dense graphs. Essentially, a graphon is a symmetric measurable function $G : I^2 \to I$, with $I := [0, 1]$ indexing a continuum of possible positions for nodes in the graph and $G(u, v)$ representing the edge density between nodes placed at u and v.

We refer to a recent series of papers by Bayraktar et al. [47, 55] for developments in the theory of graphon systems of interacting diffusions, the corresponding graphon-based limit theory, and propagation of chaos. These results are also applicable to graphon games on the underlying networks. Graphon static games have been studied in [82, 182]. For dynamic games, we refer to [101] for discrete time models and [36, 127, 162] for continuous time models. The chapter is closely related to [162], which uses the concept of graphon equilibrium to construct approximate equilibria for large finite games on any weighted, directed graph that converges in cut norm. However, unlike our work, [162] does not consider direct interactions in the dynamics, and the inhomogeneous interactions are only present in the reward function.

This chapter aims to develop a graphon interacting model to solve graphon games with heterogeneous interactions and jumps, while maintaining tractability comparable to traditional MFG. The traditional MFG framework is based on a fixed point problem describing the law of the state process $(X(t))_{t\in[0,T]}$ of a typical player. In the graphon game model, we consider a fixed point problem for a family of laws $(X_u(.))_{u\in I}$, which can be also viewed as the joint law of (U, X), where X is the state process and the uniform random variable U in $I := [0, 1]$ is interpreted as the "label" (order of vertex on network in limiting sense) of the player in the graphon. Despite the heterogeneous interactions, we also include jumps in the dynamics to model the instantaneous impacts. The jumps are induced by Poisson random measures with different intensity measures for different labels, which is a source of individual heterogeneity.

The graphon mean field model with jumps is quite useful in many fields, especially in finance and biology. For instance, consider a financial network consisting of banks or investment firms with internal links and external investments. The internal links within the network, such as shared liabilities, credit exposures, or interbank lending, could be represented by the graphon interaction, while the external investments made by each entity introduce outside risks. These external risks could be influenced by various factors, such as market fluctuations or global events, and are modeled by the Poisson random measures. This way, the model captures the complex interactions and risk exposures that characterize real-world financial networks. In addition, we incorporate control "intensively". The control term is present not only in the drift, as in [82, 162], but also in the diffusion and jump terms. Furthermore, we also have the graphon interaction in the diffusion term, which is not present in the model in [82, 162]. Combined with jumps and controls, more heterogeneity is introduced into our setup, and the interacting dynamic system becomes more complex compared to [47, 55]. Thus, the analysis becomes more involved when we try to construct the connection between finite games and graphon games. Chapter 5 provides a systemic study of graphon mean field BSDE with jumps and the associated limit theory. Although it is a system in backward form, some results on propagation of chaos can be useful here for the analysis of our graphon games.

Working directly with a continuum of players, driven by a continuum of independent Brownian motions and independent Poisson random measures, raises significant technical difficulties since neither the map $I \ni u \mapsto W_u$ nor $I \ni u \mapsto N_u$ is measurable. Noting that the value function is determined by the law of state processes $\mathcal{L}(X_u), u \in I$, we handle this issue by arguing that the laws $\mathcal{L}(X_u)$ depend measurably on u, similarly to [47, 55] but extended to a jump framework. The different intensity measures of the jump processes also increase the difficulty for measurability. We employ the canonical coupling that is present in Chapter 5, which allows us to obtain measurability in a stronger topology for the state processes X_u. Such a coupling has no influence on the graphon game, and leads to a straightforward way to investigate the connection between graphon games and finite games.

The chapter is structured as follows. In Section 6.2, we introduce the probabilistic set-up, necessary notations, and background for graphons. Section 6.3 is devoted to the main results on graphon game models with jumps and the associated graphon equilibrium issues. In Section 6.4, we study large finite networked games with heterogeneous interactions and their limit characteristics when the interaction matrix converges to a given graphon. In Section 6.5, we investigate the approximate Nash equilibria of finite games. The proofs of the main results are presented in Section 6.6.

6.2 Probabilistic set-up and notations

We introduce the probabilistic setting where we work and some notations in this section. For the knowledges on graphons and Wasserstein distance, we refer to §5.2.2 in Chapter 5. Let $T > 0$ be a fixed time horizon. Given a Polish space $\mathcal{S}$, denote by $\mathcal{D}([0,T],\mathcal{S})$ the space of RCLL (right continuous with left limits) functions from $[0,T]$ to $\mathcal{S}$, equipped with the topology of uniform convergence. Let $\mathcal{D} := \mathcal{D}([0,T],\mathbb{R})$. Denote by $\mathcal{P}(\mathcal{S})$ the space of probability measures on $\mathcal{S}$ and $\mathcal{M}_+(\mathcal{S})$ the set of nonnegative Borel measurable measures on $\mathcal{S}$. For a random variable X, $\mathcal{L}(X)$ denotes the law of X. Denote Unif$[0,1]$ the uniform measure on $[0,1]$ and further denote $\mathcal{P}_{\text{Unif}}([0,1] \times \mathcal{S})$ the set of Borel probability measures on $[0,1] \times \mathcal{S}$ with uniform first marginal. Denote by $\mathcal{M}^+_{\text{Unif}}([0,1] \times \mathcal{S})$ the set of nonnegative Borel measures on $[0,1] \times \mathcal{S}$ with uniform first marginal. We equip all spaces of measure with the topology of weak convergence. For a sequence $\{X_n\}_{n\in\mathbb{N}}$ of real-valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence of real numbers $\{a_n\}_{n\in\mathbb{N}}$, we write $X_n = O_p(a_n)$ if $\mathbb{P}(|X_n| \leqslant C|a_n|) \to 1$ as $n \to \infty$ for some constant C, and write $a_n = o(1)$ if $a_n \to 0$ as $n \to \infty$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $I = [0,1]$ and $\{W_u : u \in I\}$ be a family of i.i.d. Brownian motions defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\{N_u(dt, de) : u \in I\}$ be a family of independent Poisson measures defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with compensator $\nu_u(de)dt$ such that ν_u is a σ-finite measure on $E := \mathbb{R}_*$, with $\mathbb{R}_* := \mathbb{R}\backslash\{0\}$, equipped with its Borelian σ-algebra $\mathcal{B}(E)$, for each $u \in I$. Let $\{\widetilde{N}_u(dt, de) : u \in I\}$ be their compensator processes. Let $\mathbb{F} = \{\mathcal{F}_t, t \geqslant 0\}$ be the natural filtration associated with $\{W_u : u \in I\}$ and $\{N_u(dt, de) : u \in I\}$.

Let $T > 0$ be a fixed time horizon. Denote by P the predictable σ algebra on $[0,T] \times \Omega$.

We use the following notation: $L^2(\mathcal{F}_t)$ denotes the set of all $\mathcal{F}_t$-measurable and square integrable random variables, for $t \in [0,T]$; $\mathbb{S}^2_T$ denotes the set of real-valued RCLL adapted processes ϕ with

$$\|\phi\|_{\mathbb{S}^2_T} := \Big(\mathbb{E}\Big[\sup_{t\in[0,T]} |\phi_t|^2\Big]\Big)^{1/2} < \infty;$$

$\mathbb{H}^2_T$ denotes the set of real-valued adapted processes ϕ with

$$\|\phi\|_{\mathbb{H}^2_T} := \Big(\mathbb{E}\Big[\int_0^T |\phi_t|^2 dt\Big]\Big)^{1/2} < \infty;$$

and $\mathcal{M}\mathbb{S}^2_T$ denotes the set of all measurable functions X from I to $\mathbb{S}^2_T$: $u \mapsto X_u$, satisfying

$$\sup_{u\in I} \|X_u\|^2_{\mathbb{S}^2_T} = \sup_{u\in I} \mathbb{E}\Big[\sup_{t\in[0,T]} |X_u(t)|^2\Big] < \infty.$$

For $X \in \mathcal{M}\mathbb{S}^2_T$, we define the norm

$$\|X\|^I_{\mathbb{S}^2_T} := \sup_{u\in I} (\mathbb{E}\Big[\sup_{t\in[0,T]} |X_u(t)|^2\Big]\Big)^{1/2}.$$

We define $\mathcal{M}L^2(\mathcal{F}_t)$, and $\mathcal{M}\mathbb{H}^2_T$ similarly.

With a given metric space $\mathcal{S}$, we define the measure-valued function $\Lambda\mu : [0,1] \to \mathcal{M}_+(\mathcal{S})$ for any $\mu \in \mathcal{M}^+_{\text{Unif}}([0,1] \times \mathcal{S})$ as follows:

$$\Lambda\mu(u) := \int_{[0,1]\times\mathcal{S}} G(u,v)\delta_x \mu(dv, dx), \tag{6.1}$$

where δ_x denotes the Dirac measure concentrated at x.

For any bounded measurable function $\phi : \mathcal{S} \to \mathbb{R}$, the usual inner product is defined by

$$\langle \Lambda\mu(u), \phi \rangle = \int_{[0,1]\times\mathcal{S}} G(u,v)\phi(x)\mu(dv, dx).$$

Further, for two family of probability measures $\mu = \{\mu_u\}_{u\in I}$ and $\nu = \{\nu_u\}_{u\in I}$, we set

$$\mathcal{W}_2^{\mathcal{M}}(\mu,\nu) := \sup_{u\in I} \mathcal{W}_2(\mu_u,\nu_u), \text{ for } \mu,\nu \in \mathcal{P}(\mathcal{M}L^2(\mathcal{F}_t)) \text{ for all } t\in[0,T],$$
$$\mathcal{W}_{2,T}^{\mathcal{M}}(\mu,\nu) := \sup_{u\in I} W_{2,T}(\mu_u,\nu_u), \text{ for } \mu,\nu \in \mathcal{P}(\mathcal{M}\mathbb{S}_T^2),$$

and note that

$$\mathcal{W}_{2,T}^{\mathcal{M}}(\mu,\nu)) \geqslant \sup_{u\in I}\sup_{f} \Big| \int_{\mathbb{R}} f(x)\mu_{u,t}(dx) - \int_{\mathbb{R}} f(x)\nu_{u,t}(dx)\Big|, \quad \mu,\nu\in\mathcal{M}\mathbb{S}_T^2, \tag{6.2}$$

where the supremum is taken over all Lipschitz continuous functions $f : \mathbb{R} \to \mathbb{R}$ with Lipschitz constant 1 such that the integral exists.

6.3 Graphon mean field games with jumps

This section is dedicated to the main results on graphon games model with jumps and associated graphon equilibrium issues.

An admissible control rule is an adapted control process $\alpha_I := (\alpha_u(t))_{t\in[0,T],u\in I} \in \mathcal{M}\mathbb{H}^2$. We restrict ourselves to Markovian feedback controls. Let $\mathcal{A}_I$ be the set of graphon controls α defined as measurable function $\alpha : [0,T]\times I\times\mathbb{R} \to A; (t,u,x) \mapsto \alpha(t,u,x)$, where A is a compact metric space. We restrict ourselves to the one-dimensional case, but our results can be generalized to the multi-dimensional set-up.

The dynamics of the controlled graphon system is as follows,

$$\begin{aligned} dX_u^\alpha(s) &= \int_I\int_{\mathbb{R}} G(u,v)b(s,X_u^\alpha(s),x,\alpha(s,u,X_u^\alpha(s)))\mu_{v,t}^\alpha(dx)dvds \\ &\quad + \int_I\int_{\mathbb{R}} G(u,v)\sigma(s,X_u^\alpha(s),x,\alpha(s,u,X_u^\alpha(s)))\mu_{v,t}^\alpha(dx)dvdW_u(s) \\ &\quad + \int_E \ell(s,X_u^\alpha(s),e,\alpha(s,u,X_u^\alpha(s)))\widetilde{N}_u(ds,de), \qquad X_u(0)=\xi_u, \qquad u\in I, \end{aligned} \tag{6.3}$$

where $\mu_v^\alpha := \mathcal{L}(X_v^\alpha) \in \mathcal{P}(\mathcal{D})$ and $\mu_{v,s}^\alpha := \mathcal{L}(X_v^\alpha(s)) \in \mathcal{P}(\mathbb{R})$. We assume that $\boldsymbol{\xi} := \{\xi_u\}_{u\in I} \in \mathcal{M}L^2(\mathcal{F}_0)$, that is for each $u\in I$, $\xi_u \in L^2(\mathcal{F}_T)$ and the map $u\mapsto\xi_u$ is measurable. The coefficients $b : [0,T]\times\mathbb{R}\times\mathbb{R}\times A\to\mathbb{R}$, $\sigma : [0,T]\times\mathbb{R}\times\mathbb{R}\times A\to\mathbb{R}$ and $\ell : [0,T]\times\mathbb{R}\times E\times A\to\mathbb{R}$ are Lipschitz continuous with respect to all parameters except t. We also assume that σ^2 is bounded from below from 0. We assume moreover the following.

Assumption 6.1.

- *For each* $(t,x,u,\mu) \in [0,T]\times\mathbb{R}\times[0,1]\times\mathcal{M}^+_{\text{Unif}}([0,1]\times\mathcal{D})$*, there exists* $e\in E$ *such that the set* $K_e[\mu](t,x,u) :=$
  $\big\{\big(b(t,x,\Lambda\mu_t(u),a),\sigma^2(t,x,\Lambda\mu_t(u),a),\ell(t,x,e,a),z\big) : a\in A, z\leqslant f(t,x,\Lambda\mu_t(u),a)\big\}$ *is convex.*

- *The map $e \mapsto \ell(t,x,e,a)$ is affine for each $(t,x,a) \in [0,T] \times \mathbb{R} \times A$.*

Remark 6.1. *From the second point of Assumption 6.1, it follows that for each fixed $(t,x) \in [0,T] \times \mathbb{R}$ and any two different $e_1, e_2 \in E$, the functions $a \mapsto \ell(t,x,e_1,a)$ and $a \mapsto \ell(t,x,e_2,a)$ have the same convexity (i.e. the same shape of curvature). This, in turn, implies that if $K_e[\mu](t,x,u)$ is convex for some $e \in E$, then $K_e[\mu](t,x,u)$ is convex for all $e \in E$.*

Using the definition of $\Lambda\mu$ given in (6.1), (6.3) can be rewritten in a more general form

$$\begin{aligned} dX_u^\alpha(s) = &\int_{\mathbb{R}} b(s, X_u^\alpha(s), x, \alpha(s,u,X_u^\alpha(s)))\Lambda\mu_s^\alpha(u)(dx)ds \\ &+ \int_{\mathbb{R}} \sigma(s, X_u^\alpha(s), x, \alpha(s,u,X_u^\alpha(s)))\Lambda\mu_s^\alpha(u)(dx)dW_u(s) \\ &+ \int_E \ell(s, X_u^\alpha(s), e, \alpha(s,u,X_u^\alpha(s)))\widetilde{N}_u(ds,de), \qquad X_u(0) = \xi_u \qquad u \in I, \end{aligned} \tag{6.4}$$

where $\mu^\alpha := \mathcal{L}(X^\alpha) \in \mathcal{P}_{\text{Unif}}([0,1] \times \mathcal{D})$ and $\mu_s^\alpha := \mathcal{L}(X_s^\alpha) \in \mathcal{P}_{\text{Unif}}([0,1] \times \mathbb{R})$. When the context is clear or in the proofs, we omit the upscript of the control for notation simplicity.

For any fixed distribution $\mu \in \mathcal{P}_{\text{Unif}}([0,1] \times \mathcal{D})$ and graphon control $\alpha \in \mathcal{A}_I$, we define the following graphon objective function:

$$J_G(\mu,\alpha) := \mathbb{E}\Big[\int_I \Big(\int_0^T f(t, X_u^\alpha(t), \Lambda\mu_t(u), \alpha(t,u,X_u^\alpha(t)))dt + g(X_u^\alpha(T), \Lambda\mu_T(u))\Big)du\Big], \tag{6.5}$$

where the functions $f : [0,T] \times \mathbb{R} \times \mathcal{M}(\mathbb{R}) \times A \to \mathbb{R}$ and $g : \mathbb{R} \times \mathcal{M}(\mathbb{R}) \to \mathbb{R}$ are bounded continuous w.r.t. all parameters.

Definition 6.2 (Graphon equilibrium)**.** *A graphon equilibrium is a distribution $\mu \in \mathcal{P}_{Unif}([0,1] \times \mathcal{D})$ such that there exists $\alpha^\star \in \mathcal{A}_I$ satisfying*

$$J_G(\mu,\alpha^\star) = \sup_{\alpha \in \mathcal{A}_I} J_G(\mu,\alpha), \quad \textit{with} \quad \mu = \mathcal{L}(X^{\alpha^\star}).$$

Any $\alpha^\star$ satisfying the above is called an equilibrium control for distribution μ.

Canonical coupling and measurability. Notice that in the graphon game, the state dynamic of each label can only be influenced by the law of other labels. Thus when we couple the Brownian motions and Poisson random measures in (6.4), the law of the state for each label $\mathcal{L}(X_u)$ does not change. In addition, the graphon objective function J_G is also decided only by $\mathcal{L}(X)$. Therefore, we can study the dynamic through some coupling, under which the law of the trajectories for each label keep the same and consequently the graphon equilibrium remain the same. In order to study the state processes of the interacting system with controls, we need the measurability of $u \mapsto \mathcal{L}(X_u)$. If there is no jump included, we can simply take a common Brownian motion for all label as in [55, Lemma 2.1], but due to the presence of jumps here, we need additional care for it. To address the measurability problem, we need the same assumption on the intensity measure ν_u as in Chapter 5.

Assumption 6.2. *For each $\omega \in [1,2]$, the function $I \ni u \mapsto \Phi_u^{-1}(\omega-1) \in \mathbb{R}$ is measurable, where Φ_u denotes the cumulative distribution function of ν_u; we define $\Phi_u^{-1}(1)$ as the essential supremum and $\Phi_u^{-1}(0)$ as the essential infimum.*

Through suitable coupling called "canonical" coupling, we can obtain strong measurability of X in the space $\mathcal{M}\mathbb{S}_T^2$ and transform the original graphon system to a fully coupled system defined on the canonical space.

Define the canonical filtered probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{F}}, \bar{\mathbb{P}})$, where $\bar{\mathbb{F}} = \{\bar{\mathcal{F}}_t, t \geqslant 0\}$ and $\bar{\mathbb{P}}$ are the completed natural filtration and probability measure generated by a canonical one dimensional Brownian motion W and a Poisson random measure $N(dt, de)$ with compensator $\nu(de)dt$, where ν is uniform on $[1,2]$. Now, the canonically coupled graphon system $X_u = X_u^{\alpha}, u \in I$ is written as:

$$\begin{aligned} dX_u(s) = &\int_{\mathbb{R}} b(s, X_u(s), x, \alpha(s,u,X_u(s)))\Lambda\mu_s(u)(dx)ds \\ &+ \int_{\mathbb{R}} \sigma(s, X_u(s), x, \alpha(s,u,X_u(s)))\Lambda\mu_s(u)(dx)dW(s) \\ &+ \int_{E} \ell(s, X_u(s), \Phi_u^{-1}(e-1), \alpha(s,u,X_u(s)))\widetilde{N}(ds,de), \qquad X_u(0) = \xi_u \qquad u \in I. \end{aligned} \tag{6.6}$$

Taking advantage of the above canonical coupling, we obtain the following existence and uniqueness results for the controlled state processes, which are the solutions of the SDE system with jumps (6.4).

Lemma 6.3 (Controlled state processes)**.** *Under any control $\alpha \in \mathcal{A}_I$, there exists a unique solution X to the coupling system (6.6) such that $X \in \mathcal{M}\mathbb{S}_T^2$. Moreover, there exists a unique solution X to the original system (6.4) such that $I \ni u \mapsto \mathcal{L}(X_u)$ is measurable. Furthermore, the above assertions hold for any admissible control $\alpha_I := (\alpha_u(t))_{t\in[0,T],u\in I} \in \mathcal{M}\mathbb{H}^2$.*

Proof. For a fixed law $\mu \in \mathcal{P}_{\text{Unif}}([0,1]\times\mathcal{D})$ such that $u \mapsto \mu_u \in \mathcal{P}(\mathcal{D})$ is measurable, let us first define the map $\mu \mapsto \Phi(\mu)$ by $\Phi(\mu) := (\mathcal{L}(X_u^{\mu}) : u \in I)$, where X^{μ} satisfies (6.6) with fixed μ. By a standard contraction argument, we can prove that there exists a unique fixed point $\bar{\mu} \in \mathcal{P}_{\text{Unif}}([0,1]\times\mathcal{D})$ such that $\bar{\mu} = \Phi(\bar{\mu})$. Combining the pathwise uniqueness, we can get the uniqueness of the solution X of (6.6). We omit the details, as the proof is similar to those in [47] and in Chapter 5, despite the presence of the control term. We need to ensure the measurability of X^{μ} for each fixed measurable μ, i.e., $X^{\mu} \in \mathcal{M}\mathbb{S}_T^2$. By the preservation of measurability for the limit fixed point, we shall be able to conclude that the controlled state process X belongs to the space $\mathcal{M}\mathbb{S}_T^2$. To do so, we define the iterative equation

$$\begin{aligned} X_u^{(n)}(t) = &X_u^{(n-1)}(0) + \int_0^t \int_{\mathbb{R}} b(s, X_u^{(n-1)}(s), x, \alpha(s,u,X_u^{(n-1)}(s)))\Lambda\mu_s(u)(dx)ds \\ &+ \int_0^t \int_{\mathbb{R}} \sigma(s, X_u^{(n-1)}(s), x, \alpha(s,u,X_u^{(n-1)}(s)))\Lambda\mu_s(u)(dx)dW(s) \\ &+ \int_0^t \int_{E} \ell(s, X_u^{(n-1)}(s), \Phi_u^{-1}(e-1), \alpha(s,u,X_u^{(n-1)}(s)))\widetilde{N}(ds,de), \qquad u \in I, \end{aligned}$$

with $X^0_u(t) \equiv X_u(0)$ for $t \in [0,T]$ and all $u \in I$. Now suppose $u \mapsto X_u^{(n-1)}$ is measurable. Then $u \mapsto \alpha(\cdot, u, \cdot)$ is also measurable by its definition. By the measurability of graphon $G(u,v)$, the map $u \mapsto \int_{\mathbb{R}} b(s,x',x,a)\Lambda\mu_s(u)(dx)$ is measurable for any (s,x',a). Hence, $(u,s,x',a) \mapsto \int_{\mathbb{R}} b(s,x',x,a)\Lambda\mu_s(u)(dx)$ is measurable. Moreover since b is Lipschitz continuous, we have that $(s,u,x',a) \mapsto b(s,x',x,a)$ is measurable. Now, we have that uniformly for $(s,x) \in [0,T] \times \mathbb{R}$, $b(s,x',x,a)$ is continuous and grows at most linearly in (x',a), and the same holds for
$\int_{\mathbb{R}} b(s,x',x,a)\Lambda\mu_s(u)(dx)$ uniformly for $(s,u) \in [0,T] \times [0,1]$. It follows by [55, Lemma A.4] that

$$I \ni u \mapsto \int_0^{\cdot} \int_{\mathbb{R}} b(s, X_u^{(n-1)}(s), x, \alpha(s,u,X_u^{(n-1)}(s)))\Lambda\mu_s(u)(dx)ds \in \mathbb{S}^2_T$$

is measurable. By similar arguments, we also obtain measurability with respect to the diffusion and jumps, since they are now driven by a common Brownian motion and a common Poisson random measure. This completes the proof that $X \in \mathcal{M}\mathbb{S}^2_T$ for the coupling system (6.6), and consequently we have $u \mapsto \mathcal{L}(X_u)$ is measurable for the original graphon system (6.4). Finally, note that for any measurable control process $\alpha_I := (\alpha_u)_{u \in I} \in \mathcal{M}\mathbb{H}^2$, all the arguments above go through. $\square$

We are now ready to give our main results regarding the graphon equilibria.

Theorem 6.4 (Existence of equilibrium)**.** *There exists at least one graphon equilibrium.*

Proof. See Section 6.6.1. $\square$

Remark 6.5. *Note that [162] studies the graphon equilibrium for kernels in $L^1_+([0,1]^2)$, which are not necessarily symmetric and bounded. We can also generalize our results to this case. More details are given in the next section.*

Under some additional monotonicity assumptions adapted from the classical Lasry-Lions condition [163], we can obtain the uniqueness for the graphon equilibrium.

Theorem 6.6 (Uniqueness of equilibrium)**.** *Suppose the following monotonicity condition holds: for each $a \in A$, and any $\mu_1, \mu_2 \in \mathcal{P}_{\mathrm{Unif}}([0,1] \times \mathbb{R} \times A)$ and $t \in [0,T]$, we have*

$$\int_{[0,1]\times\mathbb{R}} \Big(g(x, \Lambda\bar{\mu}_1(u)) - g(x, \Lambda\bar{\mu}_2(u))\Big)(\bar{\mu}_1 - \bar{\mu}_2)(du,dx) < 0,$$

and

$$\int_{[0,1]\times\mathbb{R}\times A} \Big(f((t,x,\Lambda\bar{\mu}_1(u),a) - f(t,x,\Lambda\bar{\mu}_2(u),a)\Big)(\mu_1 - \mu_2)(du,dx,da) < 0,$$

where $\bar{\mu}$ is the marginal distribution of the first two coordinates. Then there exists a unique graphon equilibrium.

Proof. Suppose there are two equilibria $\bar{\mu}_1$ and $\bar{\mu}_2$. Let μ_1 be the joint distribution of $\bar{\mu}_1$ and its corresponding equilibrium control, same for μ_2. Then by the definition of equilibrium, we have

$$\int_{[0,1]\times\mathbb{R}\times A} f((t,x,\Lambda\bar{\mu}_1(u),a)\mu_1(du,dx,da) \geqslant \int_{[0,1]\times\mathbb{R}\times A} f((t,x,\Lambda\bar{\mu}_1(u),a)\mu_2(du,dx,da)$$

and

$$\int_{[0,1]\times\mathbb{R}\times A} f((t,x,\Lambda\bar{\mu}_2(u),a)\mu_2(du,dx,da) \geqslant \int_{[0,1]\times\mathbb{R}\times A} f((t,x,\Lambda\bar{\mu}_2(u),a)\mu_1(du,dx,da).$$

The above two formulas lead to

$$\int_{[0,1]\times\mathbb{R}\times A}\Big(f((t,x,\Lambda\bar{\mu}_1(u),a)-f(t,x,\Lambda\bar{\mu}_2(u),a)\Big)(\mu_1-\mu_2)(du,dx,da)\geqslant 0,$$

and the same for g, which is a contradiction to our assumptions. □

Let J_G^{u,ξ_u} be the marginal value function of graphon equilibrium defined as

$$J_G^{u,\xi_u}(\mu,\alpha) := \mathbb{E}\bigg[\int_0^T f(t,X(t),\Lambda\mu(u),\alpha(t,X(t)))dt + g(X(T),\Lambda\mu(u))\bigg], \tag{6.7}$$

with the dynamics of $X = X^\alpha$ being

$$\begin{aligned} dX(s) &= \int_{\mathbb{R}} b(s,X(s),x,\alpha(s,X(s)))\Lambda\mu_s(u)(dx)ds \\ &\quad + \int_{\mathbb{R}} \sigma(s,X(s),x,\alpha(s,X(s)))\Lambda\mu_s(u)(dx)dW(s) \\ &\quad + \int_E \ell(s,X(s),e,\alpha(s,X(s)))\widetilde{N}(ds,de), \qquad X(0)=\xi_u. \end{aligned}$$

The following proposition tell us that when the population distribution is given, the optimal Markovian feedback control for each label $u \in I$ is also the optimal control in the set $\mathcal{A}_1$ of measurable functions $[0,T]\times\mathbb{R}\to A$. Since we fix the population distribution, it can be viewed as if there were no interaction in the dynamics. This proposition is an analogue to [162, Lemma 4].

Proposition 6.7 (Marginal supremum)**.** *Given $\mu \in \mathcal{P}_{Unif}([0,1]\times\mathcal{D})$, if $\alpha^\star \in \mathcal{A}_I$ attains the supremum of $J_G(\mu,\alpha)$ defined in (6.5) over all α in $\mathcal{A}_I$, then for a.e. $u \in I$, we have*

$$J_G^{u,\xi_u}(\mu,\alpha_u^\star) = \sup_{\alpha\in\mathcal{A}_1} J_G^{u,\xi_u}(\mu,\alpha),$$

with $\alpha_u^\star(t,x) := \alpha^\star(t,u,x)$, where $J_G^{u,\xi_u}(\mu,\alpha)$ is defined in (6.7).

We omit the proof which follows the arguments of the proof of [162, Lemma 4] and can easily be adapted to our set-up.

6.4 Finite networked games with heterogenous interactions

In this section, we study large finite networked games with heterogeneous interactions and analyze their limiting characteristics as the number of players n approaches infinity, with the interaction matrix converging to a given graphon.

6.4.1 Finite game with jumps

Let $n \in \mathbb{N}$ be the size of the network. We consider an inhomogeneous interacting particle system $X^{(n)} = X^{(n),\alpha}$ with controlled dynamics

$$\begin{aligned} dX_i^{(n)}(s) =& \frac{1}{n}\sum_{j=1}^{n} \zeta_{ij}^{(n)} b(s, X_i^{(n)}(s), X_j^{(n)}(s), \alpha(s, X^{(n)}(s)))ds \\ &+ \frac{1}{n}\sum_{j=1}^{n} \zeta_{ij}^{(n)} \sigma(s, X_i^{(n)}(s), X_j^{(n)}(s), \alpha(s, X^{(n)}(s)))dW_i(s) \\ &+ \int_E \ell(s, X_i^{(n)}(s), e, \alpha(s, X^{(n)}(s)))\widetilde{N}_i(ds, de), \qquad X_i^{(n)}(0) = \xi_i^{(n)}, \end{aligned} \tag{6.8}$$

where $\{W_i, i \in [n]\}$ are i.i.d. Brownian motions, and $\{N_i(dt, de), i \in [n]\}$ are independent Poisson random measures with compensator $\nu_i(de)dt$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that ν_i is a σ-finite measure on $E := \mathbb{R}_*$, with $\mathbb{R}_* := \mathbb{R}\backslash\{0\}$, equipped with its Borelian σ-algebra $\mathcal{B}(E)$. Let $\{\widetilde{N}_i(dt, de) : u \in I\}$ be the compensator processes. Let $\mathbb{F} = \{\mathcal{F}_t, t \geqslant 0\}$ be the natural filtration associated with W_i and $\widetilde{N}_i(dt, de)$, $i = 1, \ldots, n$. We assume that $\xi_i^{(n)} \in L^2(\mathcal{F}_0)$ for all $i = 1, \ldots, n$.

Here, $\zeta^{(n)} := (\zeta_{ij}^{(n)})_{ij}$ is an $n \times n$ symmetric matrix with nonnegative entries, called the interaction matrix. It describes the strength of interaction between players i and j. Usually $\zeta^{(n)}$ is the weights between edges, but it can also be defined as the probability that an edge is present between the vertices.

The controls α are in Markovian feedback form, i.e., they are in the set $\mathcal{A}_n$ of measurable functions from $[0, T] \times \mathbb{R}^n \to A$, where A is a compact metric space representing the set of actions. We denote by $\mathcal{A}_n^n$ the set of vectors $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$, with $\alpha_i \in \mathcal{A}_n$.

We assume that the coefficients $b : [0, T] \times \mathbb{R} \times \mathbb{R} \times A \to \mathbb{R}$, $\sigma : [0, T] \times \mathbb{R} \times \mathbb{R} \times A \to \mathbb{R}$, and $\ell : [0, T] \times \mathbb{R} \times E \times A \to \mathbb{R}$ are Lipschitz continuous w.r.t. all parameters except t. We further assume that σ^2 is bounded from below.

For each player $i \in [n]$, we define the neighborhood empirical measure as

$$M_i^{(n)} := \frac{1}{n}\sum_{j=1}^{n} \zeta_{ij}^{(n)} \delta_{X_i^{(n)}} \in \mathcal{P}(\mathcal{D}), \tag{6.9}$$

and the neighborhood empirical measure at time s,

$$M_i^{(n)}(s) := \frac{1}{n}\sum_{j=1}^{n} \zeta_{ij}^{(n)} \delta_{X_i^{(n)}(s)} \in \mathcal{P}(\mathbb{R}).$$

Each player $i \in [n]$ seeks to optimize its own objective function with respect to the control $\boldsymbol{\alpha}$:

$$J_i(\boldsymbol{\alpha}) := \mathbb{E}\Big[\int_0^T f(t, X_i^{(n)}(t), M_i^{(n)}(t), \alpha(t, X^{(n)}(t)))dt + g(X_i^{(n)}(T), M_i^{(n)}(T))\Big],$$

where the functions f and g are bounded continuous w.r.t. all parameters.

6.4.2 Propagation of chaos for controlled graphon system

To approximate Nash equilibria for finite games on a network, we must establish a connection between the finite system and the graphon system, specifically the relationship between their laws of state processes. In this section, we present the results of the propagation of chaos. As previously mentioned, the canonical coupling significantly simplifies the analysis, and sometimes we study under canonical coupling.

First, we provide a result to measure the distance between the state processes induced by different graphons.

Theorem 6.8 (Stability of graphon)**.** *Let X^{α} and $X^{(n),\alpha}$ be the solutions of (6.3) associated with graphon G and G_n, initial conditions ξ and $\xi^{(n)}$ respectively and both under control α. Suppose that for each $u \in I$, $\alpha(t,u,x)$ is Lipschitz in x. Then, we have (for some constant $C > 0$)*

$$\mathbb{E}\Big[\int_I \sup_{t\in[0,T]} |X_u^{(n),\alpha}(t) - X_u^{\alpha}(t)|^2 du\Big] \leqslant C\Big(\int_I \mathbb{E}|\xi_u - \xi_u^{(n)}|^2 du + \|G - G_n\|_{\square}\Big).$$

Moreover, we have

$$\sup_{u\in I} \|X_u^{(n),\alpha}(t) - X_u^{\alpha}(t)\|_{\mathbb{S}_T^2}^2 \leqslant C\Big(\sup_{u\in I} \mathbb{E}|\xi_u^{(n)} - \xi_u|^2 + \|G - G_n\|_{\infty\to\infty}\Big).$$

Proof. See Section 6.6.2. □

The above stability result illustrates the difference between two systems. In particular, if $\|G_n - G\|_{\square} \to 0$ and $\mathbb{E}[\int_I |\xi_u - \xi_u^{(n)}|du] \to 0$ as $n \to \infty$, it follows that

$$\mathbb{E}\Big[\int_I \sup_{t\in[0,T]} |X_u^{(n),\alpha}(t) - X_u^{\alpha}(t)|^2 du\Big] \to 0.$$

Under certain continuity assumptions on the graphon and initial condition, we can achieve continuity concerning the state processes.

Assumption 6.3. *There exists a finite collection of intervals $\{I_i : i = 1, \dots, n\}$ such that $I = \bigcup_i I_i$ and, for each $i \in \{1, \dots, n\}$, we have:*

(i) $u \to \mathcal{L}(\xi_u)$ is continuous a.e. on I_i w.r.t. the $\mathcal{W}_2$ metric.

(ii) For each $j \in \{1, \dots, n\}$, $G(u,v)$ is continuous in u and v a.e. on $I_i \times I_j$.

(iii) The intensity measure ν_u is continuous in u for the Wasserstein distance $\mathcal{W}_2$ on each I_i.

Assumption 6.4. *There exists a finite collection of intervals $\{I_i : i = 1, \dots, n\}$ such that $I = \bigcup_i I_i$, and for some constant C, we have for all $u_1, u_2 \in I_i$, $v_1, v_2 \in I_j$, and $i, j \in \{1, \dots, N\}$,*

$$\mathcal{W}_2(\mathcal{L}(\xi_{u_1}), \mathcal{L}(\xi_{u_2})) \leqslant C|u_1 - u_2|, \tag{6.10}$$

$$|G(u_1, v_1) - G(u_2, v_2)| \leqslant C(|u_1 - u_2| + |v_1 - v_2|),$$

and

$$\mathcal{W}_2(\nu_{u_1}, \nu_{u_2}) \leqslant C|u_1 - u_2|.$$

We then have the following result regarding the difference between labels within the same system.

Lemma 6.9. *We have the following:*

(i) (Continuity) Suppose Assumption 6.3 holds. Furthermore, assume that for each $\{I_i, i = 1, \ldots, n\}$, the control $\alpha(t, u, x)$ is continuous in u on each I_i and continuous in x. Then the map $I_i \ni u \to \mathcal{L}(X_u^\alpha)$ is continuous w.r.t. the $\mathcal{W}_{2,T}$ distance for each I_i.

(ii) (Lipschitz continuity) Suppose Assumption 6.4 holds. Furthermore, assume that for each $\{I_i, i = 1, \ldots, n\}$, the control $\alpha(t, u, x)$ is Lipschitz continuous in u on each I_i and Lipschitz continuous in x. Then the map $I_i \ni u \to \mathcal{L}(X_u)$ is also Lipschitz continuous w.r.t. the $\mathcal{W}_{2,T}$ distance for each I_i.

Proof. See Section 6.6.3. □

To examine the relationship between the finite system and the graphon system, we require the following regularity assumption on the strength of interaction $\zeta^{(n)}$, similarly as in Chapter 5.

Assumption 6.5 (Interaction regularity)**.** *We say $\zeta^{(n)} := \{\zeta_{ij}^{(n)}\}_{i,j\in[n]}$ satisfies the regularity assumption with graphon G if either:*

(i) $\zeta_{ij}^{(n)} = G(\frac{i}{n}, \frac{j}{n})$;

(ii) $\zeta_{ij}^{(n)} = \mathrm{Bernoulli}\Big(G(\frac{i}{n}, \frac{j}{n})\Big)$ independently for all $1 \leqslant i \leqslant j \leqslant n$ and independent of $\{W_u, N_u, \xi_u : u \in I\}$ and $\{W_i, N_i, \xi_i : i \in [n]\}$.

We call $\{G_n\}_{n\in\mathbb{N}}$ a sequence of step graphonsa sequence of step graphons if, for each $n \in \mathbb{N}$, G_n is a graphon and satisfies $G_n(u, v) = G_n\Big(\frac{\lceil nu\rceil}{n}, \frac{\lceil nv\rceil}{n}\Big)$ for all $(u, v) \in I \times I$. With the above two results, we obtain the following convergence results from the finite controlled system to the graphon controlled system:

Theorem 6.10 (Large population convergence)**.** *Let $\alpha(t, u, x)$ be a Lipschitz function on (u, x), and let $\alpha_i^{(n)}(t, x) = \alpha(t, \frac{i}{n}, x)$. Let $X^{(n)}$ and X be the solutions of (6.8) and (6.3) respectively, with initial conditions $\xi^{(n)}$ and ξ, controls $\alpha^{(n)} := (\alpha_i^{(n)})_{i\in[n]}$ and α. Suppose Assumption 6.4 holds with G, and $\zeta^{(n)}$ satisfies the regularity Assumption 6.5 with G_n, where $\{G_n\}_n$ is a sequence of step graphons such that $\|G - G_n\|_\square \to 0$. Then we have the following convergence result for the empirical mean of the neighborhood measure (defined in (6.9)):*

$$\frac{1}{n}\sum_{i=1}^{n} M_i^{(n)} \to \int_I \Lambda\mu(v)dv,$$

in probability in the weak sense, where $\mu := \mathcal{L}(X)$. Furthermore, for each $i \in [n]$ and any Lipschitz continuous bounded function h from $\mathcal{D}$, we have (for some constant $C > 0$)

$$\mathbb{E}\Big[\langle h, M_i^{(n)}\rangle - \langle h, \Lambda\mu(\frac{i}{n})\rangle\Big]^2 \leqslant \frac{C}{n}\sum_{j=1}^{n}\mathbb{E}\Big|\xi_j^{(n)} - \xi_{\frac{j}{n}}\Big|^2 + C\|G_n - G\|_{\square} + C\|G_n - G\|_{\infty\to\infty} + \frac{C}{n}.$$

If W_i, N_i and $W_{\frac{i}{n}}, N_{\frac{i}{n}}$ are the same for each $i \in [n]$, then we have

$$\frac{1}{n}\sum_{i=1}^{n}\|X_i^{(n)} - X_{\frac{i}{n}}\|_{\mathbb{S}_T^2}^2 \leqslant C\Big(\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_{\square} + \frac{1}{n}\Big),$$

and moreover

$$\max_{i\in[n]}\|X_i^{(n)} - X_{\frac{i}{n}}\|_{\mathbb{S}_T^2}^2 \leqslant C\Big(\max_{i\in[n]}\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_{\infty\to\infty} + \frac{1}{n}\Big).$$

Proof. See Section 6.6.4. □

Remark 6.11. *Notice that for any $t \in [0, T]$, $\mathbb{S}_T^2 \ni X \mapsto X_t \in \mathbb{R}$ is continuous. Thus, for any bounded Lipschitz continuous function $H : \mathbb{R} \to \mathbb{R}$, we have that $\langle H, M_i^{(n)}(t)\rangle \to \langle H, \Lambda\mu_t(\frac{i}{n})\rangle$ with the same convergence rate as $\langle h, M_i^{(n)}\rangle \to \langle h, \Lambda\mu(\frac{i}{n})\rangle$, under the conditions in Theorem 6.10.*

When the graphon is not necessarily continuous, we can still obtain similar convergence results. To this end, we introduce the following definition.

Definition 6.12 (Continuous modification set)**.** *Let (X, d_x) and (Y, d_y) be two metric spaces, and let $f : X \to Y$ be measurable. We say that a point $x \in X$ is in the continuous modification set of f if x belongs to some $A \in \mathcal{B}(X)$ such that the restriction of f on A is continuous.*

As a corollary of Theorem 6.10, we have the following convergence result for general graphons.

Corollary 6.13 (Convergence for general graphon)**.** *Let $\alpha(t, u, x)$ be a Lipchitz function in (u, x), and let $\alpha_i^{(n)}(t, x) = \alpha(t, \frac{i}{n}, x)$. Let $X^{(n)}$ and X be the solutions of (6.8) and (6.3), respectively, with initial conditions $\xi^{(n)}$ and ξ, controls $\alpha^{(n)}$ and α. Suppose that for each $n \in \mathbb{N}$, $\{(\frac{i}{n}, \frac{j}{n}), i, j = 1, \ldots, n\}$ is in the continuous modification set of G. Suppose Assumption 6.3 (a) holds, $\zeta^{(n)}$ satisfies the regularity Assumption 6.5 with G_n, and $\{G_n\}_n$ is a sequence of step graphons such that $\|G - G_n\|_{\square} \to 0$, for any G. Then we have*

$$\frac{1}{n}\sum_{i=1}^{n} M_i^{(n)} \to \int_I \Lambda\mu(v)dv,$$

in probability in the weak sense, where $\mu := \mathcal{L}(X)$. Furthermore, if $\|G_n - G\|_{\infty\to\infty} \to 0$, then for all $i \in [n]$, $M_i^{(n)} \to \Lambda\mu(\frac{i}{n})$. If W_i, N_i and $W_{\frac{i}{n}}, N_{\frac{i}{n}}$ are the same for each $i \in [n]$, then we have

$$\frac{1}{n}\sum_{i=1}^{n}\|X_i^{(n)} - X_{\frac{i}{n}}\|_{\mathbb{S}_T^2}^2 \to 0,$$

and if $\|G_n - G\|_{\infty\to\infty} \to 0$, then $\max_{i\in[n]}\|X_i^{(n)} - X_{\frac{i}{n}}\|_{\mathbb{S}_T^2}^2 \to 0$.

Proof. See Section 6.6.5 □

We end this section by the following example of a graphon which is nowhere continuous but satisfy the conditions of Corollary 6.13.

Example 6.14 (Dirichlet graphon)**.** *Consider the graphon G defined by $G(u,v)=1$ if $u,v\in\mathbb{Q}\cap[0,1]$, and $G(u,v)=0$ otherwise. Although G is nowhere continuous, all rational points belong to the continuous modification set of G, and thus the results of Corollary 6.13 apply to this graphon.*

6.5 Approximate Nash equilibria of finite games

For $\boldsymbol{\epsilon}=(\epsilon_1,\dots,\epsilon_n)\in[0,\infty)^n$, the $\boldsymbol{\epsilon}$-Nash equilibrium is defined as any $\boldsymbol{\alpha}=(\alpha_1,,\dots,\alpha_n)\in\mathcal{A}_n^n$ satisfying for all $i\in[n]$,

$$J_i(\boldsymbol{\alpha})\geqslant\sup_{\beta\in\mathcal{A}_n}J_i(\alpha_1,\dots,\alpha_{i-1},\beta,\alpha_{i+1},\dots,\alpha_n)-\epsilon_i.$$

For any $\epsilon>0$, a global ϵ-Nash equilibrium is defined as any $\boldsymbol{\alpha}\in\mathcal{A}_n^n$ satisfying for all $i\in[n]$,

$$J_i(\boldsymbol{\alpha})\geqslant\sup_{\beta\in\mathcal{A}_n}J_i(\alpha_1,\dots,\alpha_{i-1},\beta,\alpha_{i+1},\dots,\alpha_n)-\epsilon.$$

We use the equilibrium control for graphon games as a benchmark to infer the equilibrium for finite games. With the propagation of chaos results, as the population size grows, the distributions of state processes of finite games and graphon games become closer. Intuitively, the equilibrium control for each player in the finite game should be very close to that taken for the corresponding label in the limit graphon system. It is natural to choose the control associated to label $\frac{i}{n}$ for the i-th player in an n-player game. When the graphon equilibrium control has some continuity with respect to u, we could just consider controls associated with labels close to $\frac{i}{n}$.

Let us define

$$\epsilon_i^{(n)}(\boldsymbol{u}^{(n)}):=\sup_{\beta\in\mathcal{A}_n}J_i(\alpha^\star(u_1^{(n)}),\dots,\alpha^\star(u_{i-1}^{(n)}),\beta,\alpha^\star(u_{i+1}^{(n)})\dots,\alpha^\star(u_n^{(n)}))-J_i(\boldsymbol{\alpha}^\star),\tag{6.11}$$

where $\boldsymbol{u}^{(n)}:=(u_1^{(n)},\cdots,u_n^{(n)})$, and $u_i^{(n)}$ is such that $\alpha^\star(u_i^{(n)}):=\alpha^\star(\cdot,u_i^{(n)},\cdot)$, i.e., player i uses the control rule of the graphon equilibrium control of label $u_i^{(n)}$.

The accuracy and complexity of the approximate equilibria for finite games depend on the underlying graphon and on the way in which the network converges to its graphon.

We start with the simplest case of a piecewise constant graphon.

Piecewise constant graphon. We call a graphon *piecewise constant* if there exists a collection of intervals $\{I_i,i=1,\dots,k\}$ for some $k\in\mathbb{N}$ such that $I=\bigcup_{i=1}^k I_i$ and for all $u_1,u_2\in I_i$, $v_1,v_2\in I_j$, and $i,j\in\{1,\dots,k\}$, we have $G(u_1,v_1)=G(u_2,v_2)$ and $G(u_1,v_2)=G(u_2,v_1)$. Such a graphon corresponds to the *stochastic block model* and can be thought of as a model of multi-type mean field games.

Theorem 6.15 (Piecewise constant graphon)**.** *Suppose G is a piecewise constant graphon and $\zeta^{(n)}$ satisfies regularity Assumption 6.5 with G. Suppose also that Assumption 6.3 (i) holds. If*

$$\max_{i=1,\ldots,n} \mathbb{E}\big|\xi_i^{(n)} - \xi_{\frac{i}{n}}\big|^2 = O(n^{-1}),$$

then taking $u_i^{(n)} = \frac{i}{n}$, we have as $n \to \infty$,

$$\max_{i=1,\ldots,n} \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \to 0.$$

Moreover, if the initial condition is Lipschitz, satisfying (6.10), then we have

$$\max_{i=1,\ldots,n} \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) = O(n^{-1}).$$

Proof. We omit the proof since it follows the same arguments as the proof of the Lipschitz case below, but are simpler. ☐

(Lipschitz) Continuous graphon. We call a graphon $G(u, v)$ *continuous* if there exists a collection of intervals $\{I_i, i = 1, \ldots, k\}$, for some $k \in \mathbb{N}$, such that $I = \bigcup_i I_i$, and G is piecewise continuous with respect to u and v in all intervals $I_i, i = 1, \ldots, k$. Furthermore, we call it *Lipschitz continuous* if for all $u_1, u_2 \in I_i$, $v_1, v_2 \in I_j$, and $i, j \in 1, \ldots, k$, there exists a constant C such that

$$|G(u_1, v_1) - G(u_2, v_2)| \leqslant C(|u_1 - u_2| + |v_1 - v_2|).$$

To study continuous graphons, we need to assume continuity of graphon equilibrium control with respect to the label. We introduce the following concavity assumption.

Assumption 6.6 (Concavity)**.**

- *$f(t, x, \mu, a)$ is concave in x and strictly concave in a.*
- *For all $\lambda \in [0, 1]$, $a_1, a_2 \in A$,*

$$\lambda f(t, x, \mu, a_1) + (1 - \lambda) f(t, x, \mu, a_2) \leqslant f(t, x, \mu, \bar{a}_\lambda),$$

where $\bar{a}_\lambda = \bar{a}_\lambda(t, x, \mu)$ is the solution to

$$b(t, x, \mu, \bar{a}_\lambda) = \lambda b(t, x, \mu, a_1) + (1 - \lambda) b(t, x, \mu, a_2).$$

Remark 6.16. *This assumption is satisfied in particular when the drift b is affine in a, as assumed in [162, Theorem 3].*

Under the above assumption, we can obtain the following stability lemma for graphon equilibrium control.

Lemma 6.17 (Stability of control)**.** *Suppose that Assumptions 6.3 and 6.6 are satisfied. Then there exists a unique optimizer $\alpha_u^\star$ for $\sup_{\alpha\in\mathcal{A}_n} J_G^{u,\xi_u}(\mu,\alpha)$. Let $\alpha^\star(t,u,x) := \alpha_u^\star(t,x)$. We have $\alpha^\star(t,u,x)$ is (piecewise) continuous in (u,x), and the law of $X_u^{\alpha^\star}$ is (piecewise) continuous in u in the weak sense. Furthermore, if G, f, g are all Lipschitz continuous and Assumption 6.4 is satisfied, then all the continuities become Lipschitz continuities.*

Proof. See Section 6.6.6. □

Using this lemma, we obtain the following approximate result. For each $i \in [n]$, we define $\mathcal{I}_i^{(n)} := (\partial_- I_j, \frac{i}{n}]$ if $\frac{i-1}{n} \notin I_j$, $\frac{i}{n} \in I_j$; $\mathcal{I}_i^{(n)} := \left(\frac{i-1}{n}, \frac{i}{n}\right]$ if $\frac{i-1}{n}, \frac{i}{n} \in I_j$; $\mathcal{I}_i^{(n)} := \left[\frac{i-1}{n}, \partial_+ I_j\right)$ if $\frac{i}{n} \in I_j$ and $\frac{i+1}{n} \notin I_j$, where ∂_- and ∂_+ denote the lower and upper borders, respectively.

Theorem 6.18. *Suppose Assumption 6.6 holds, $\zeta^{(n)}$ satisfies the regularity Assumption 6.5 with step graphon G_n, and $\|G - G_n\|_\square \to 0$.*

(i) (Continuous graphon). Suppose Assumption 6.3 holds, G is continuous, and the initial condition satisfies $\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 \to 0$. Then we have

$$\text{ess sup}_{\boldsymbol{u}^{(n)}\in\mathcal{I}_1^{(n)}\times\cdots\times\mathcal{I}_n^{(n)}} \frac{1}{n}\sum_{i=1}^n \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \to 0.$$

Furthermore, if $\|G - G_n\|_{\infty\to\infty} \to 0$ and $\max_{i=1,\dots,n} \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 \to 0$, then we have

$$\text{ess sup}_{\boldsymbol{u}^{(n)}\in\mathcal{I}_1^{(n)}\times\cdots\times\mathcal{I}_n^{(n)}} \max_{i=1,\dots,n} \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \to 0.$$

(ii) (Lipschitz Continuous graphon). Suppose Assumption 6.4 holds, G, f, and g are Lipschitz continuous, and the initial condition satisfies $\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 = O(n^{-1})$. Then we have

$$\text{ess sup}_{\boldsymbol{u}^{(n)}\in\mathcal{I}_1^{(n)}\times\cdots\times\mathcal{I}_n^{(n)}} \frac{1}{n}\sum_{i=1}^n \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) = O(n^{-1}).$$

Furthermore, if $\|G - G_n\|_{\infty\to\infty} \to 0$ and $\max_{i=1,\dots,n} \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 = O(n^{-1})$, then we have

$$\text{ess sup}_{\boldsymbol{u}^{(n)}\in\mathcal{I}_1^{(n)}\times\cdots\times\mathcal{I}_n^{(n)}} \max_{i=1,\dots,n} \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) = O(n^{-1}).$$

Proof. See Section 6.6.7 □

For a non-continuous graphon, we obtain a slightly weaker result.

Proposition 6.19 (General graphon)**.** *Suppose Assumption 6.6 holds and for each $n \in \mathbb{N}$, $\{(\frac{i}{n}, \frac{j}{n})i, j \in [n]\}$ is in the continuous modification set of G (see Definition 6.12). Suppose moreover that $\zeta^{(n)}$ satisfies the regularity Assumption 6.5 with step graphon G_n, and $\|G - G_n\|_\Box \to 0$. If the initial condition satisfies $\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 \to 0$, then taking $u_i^{(n)} = \frac{i}{n}$, we have, as $n \to \infty$,*

$$\frac{1}{n}\sum_{i=1}^n \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \to 0.$$

Proof. The proof follows, similarly to the proof of Theorem 6.18, using Corollary 6.13. □

Sampling graphon. Let $U_1, \dots, U_n$ be i.i.d. uniform random variables on $[0, 1]$, and let $U_{(1)}, \dots, U_{(n)}$ be their order statistics. We say that $\zeta^{(n)}$ is *sampled with weights* from the graphon G if $\zeta_{ij}^{(n)} = G(U_{(i)}, U_{(j)})$. We say that $\zeta^{(n)}$ is *sampled with probabilities* from the graphon G if

$$\zeta_{ij}^{(n)} = \text{Bernoulli}\big(G(U_{(i)}, U_{(j)})\big).$$

It is clear that if the strength of interactions $\zeta^{(n)}$ is sampled from a graphon, it will introduce more randomness into the system and, hence, hinder our analysis. However, as the number of players n becomes very large, the randomness can be reduced and does not interfere with the approximation for equilibrium.

Theorem 6.20 (Sampling graphon)**.** *Suppose Assumption 6.6 and 6.3 hold. Let $\zeta^{(n)}$ be sampled from the continuous graphon G. If the initial condition satisfies $\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 \to 0$, then we have, for both ways of sampling above, as $n \to \infty$,*

$$\text{ess sup}_{\boldsymbol{u}^{(n)} \in \mathcal{I}_1^{(n)} \times \dots \times \mathcal{I}_n^{(n)}} \frac{1}{n}\sum_{i=1}^n \epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \to 0.$$

Proof. See Section 6.6.7. □

6.6 Proofs

6.6.1 Existence of graphon equilibrium: Proof of Theorem 6.4

We first prove the existence of a relaxed equilibrium and then show how to construct a strict control.

A relaxed control rule is a measure on $[0, T] \times A$ with the first marginal equal to the Lebesgue measure. Denote $\mathcal{V}$ as the set of relaxed controls. Since A is assumed to be compact, $\mathcal{V}$ is also compact, equipped with the weak topology. For each $m \in \mathcal{V}$, we have $m(dt, da) = dtm_t(da)$ with m_t measurable and unique up to almost everywhere (a.e.) equality.

A strict control is a specific control rule which involves no measure on A, that is of the form $m_t = \delta_{a(t)}$ for a.e. t with measurable $a : [0, T] \mapsto A$. Under some appropriate conditions, we will show how to construct strict controls based on relaxed controls.

Existence of a relaxed equilibrium

For all $\phi \in C_c^\infty(\mathbb{R})$ and each $u \in I$, let $\mathcal{H}_u$ be an integro-differential operator associated with label u of the following form

$$\begin{aligned}\mathcal{H}_u\phi(t,x,\mu,a) :=& b(t,x,\mu,a)\frac{\partial \phi}{\partial x}(x) + \frac{\sigma^2}{2}(t,x,\mu,a)\frac{\partial^2\phi}{\partial x^2}(x) \\ &+ \int_{\mathbb{R}}\Big(\phi(x+\ell(t,x,e,a)) - \phi(x) - \ell(t,x,e,a)\frac{\partial \phi}{\partial x}(x)\Big)\nu_u(de).\end{aligned}$$

Definition 6.21 (Controlled graphon martingale problem)**.** *For a fixed $\mu \in \mathcal{P}_{\mathrm{Unif}}([0,1]\times\mathcal{D})$ and $\lambda_0 \in \mathcal{P}_{\mathrm{Unif}}([0,1]\times\mathbb{R})$, let $\mathcal{R}(\mu)$ be the set of $\mathbb{P} \in \mathcal{P}(\Omega_m)$, $\Omega_m := \mathcal{V}\times[0,1]\times\mathcal{D}$ such that:*

(i) $\mathbb{P}\circ(U,X(0))^{-1} = \lambda_0$*;*

(ii) for each $\phi \in C_c^\infty(\mathbb{R})$ (the set of C^∞ functions with compact support), the following process $(M_t^{\mu,\phi})_{t\in[0,T]}$ is a $\mathbb{P}$-martingale,

$$M_t^{\mu,\phi}(m,u,x) := \phi(x_t) - \int_{[0,t]\times A}\mathcal{H}_u\phi(s,x_s,\Lambda\mu_s(u),a)m(ds,da), \qquad t\in[0,T].$$

Let $\Gamma^\mu(m,u,x) : \mathcal{R}(\mu) \mapsto \mathbb{R}$ be defined as

$$\Gamma^\mu(m,u,x) := \int_{[0,T]\times A} f(t,x,\Lambda\mu_t(u),a)m_u(dt,da) + g(x,\Lambda\mu_T(u)).$$

Let $\mathcal{R}_0$ be the set of $\mathbb{P}\in\mathcal{P}_{\mathrm{Unif}}([0,1]\times\mathcal{D})$ such that:

(i) $\mathbb{P}\circ(U,X(0))^{-1} = \lambda_0$.

(ii) For some constant M, $\sup_{u\in I}\mathbb{E}\|X\|^2_{\mathbb{S}^2_T} \leqslant M$.

(iii) For each nonnegative $\phi\in C_c^\infty(\mathbb{R})$, $\phi(X_t) + C(\phi)t$ is a $\mathbb{P}$-submartingale, where $C(\phi)$ is the supremum of $|\mathcal{H}_u\phi|$ over $[0,1]\times[0,T]\times\mathbb{R}\times\mathcal{M}_+(\mathbb{R})\times A$.

It is easy to see that $\mathcal{R}_0$ is nonempty, convex and closed. The first marginal is compact. Since b,σ and ℓ are bounded, by the tightness criterion of Stroock and Varadhan [190, Theorem 1.4.6], (i) and (iii) imply that the third marginal of $\mathcal{R}_0$ is tight. Condition (ii) guarantees it is relatively compact in $\mathcal{P}(\mathcal{D})$, see [198, Theorem 7.12]. Thus $\mathcal{R}_0$ is compact.

With $\mu\in\mathcal{P}_{\mathrm{Unif}}([0,1]\times\mathcal{D})$ given, define the set

$$\mathcal{R}^\star(\mu) := \arg\max_{\mathbb{P}\in\mathcal{R}(\mu)}\langle\Gamma^\mu,\mathbb{P}\rangle.$$

Define further the set-valued map $\Phi(\mu) : \mathcal{P}_{\mathrm{Unif}}([0,1]\times\mathcal{D}) \mapsto 2^{\mathcal{P}_{\mathrm{Unif}}([0,1]\times\mathcal{D})}$ by

$$\Phi(\mu) := \{\mathbb{P}\circ(X,U)^{-1} : \mathbb{P}\in\mathcal{R}^\star(\mu)\}.$$

We proceed to prove the existence by using Kakutani-Fan-Glicksberg fixed point theorem for set-valued functions. This is a powerful tool for equilibrium analysis in MFGs, as shown in previous works such as [84, 160, 162]. While the basic logic and procedures are similar to those in standard MFGs, there are additional considerations needed due to the presence of the label variable U in our model. We refer to [6, Chapter 17] for the necessary knowledge on set-valued analysis.

We introduce the definition of the continuity of set-valued functions and some basic results. For two metric spaces A and B, a set valued function $f : A \to 2^B$ is lower hemicontinuous if, whenever $x_n \to x$ in A and $y \in f(x)$, there exists $y_{n_k} \in f(x_{n_k})$ such that $y_{n_k} \to y$. If $f(x)$ is closed for each $x \in A$, we say that f is upper hemicontinuous if whenever $x_n \to x$ and $y_n \in f(x_n)$ for each n, the sequence (y_n) has a limit point in $f(x)$. Moreover, f is said continuous if it is both upper and lower hemicontinuous. If B is compact, then the graph $\{(x, y) : x \in A, y \in f(x)\}$ of f is closed if and only if $f(x)$ is closed for each $x \in A$ and f is upper hemicontinuous. To prove that the set of solutions of the set-valued fixed point equation $\mu = \Phi(\mu)$ is nonempty, we have to verify the following three conditions:

(i) $\Phi(\mu) \subset \mathcal{R}_0$ for each $\mu \in \mathcal{R}_0$.

(ii) $\Phi(\mu)$ is nonempty and convex for each $\mu \in \mathcal{R}_0$.

(iii) the graph $\{(\mu, \mu') : \mu \in \mathcal{R}_0, \mu' \in \Phi(\mu)\}$ is closed.

For any $\mathbb{P} \in \mathcal{R}_0$, let $\mu = \mathbb{P} \circ (U, X)^{-1}$. Clearly the image set $\Phi(\mu)$ satisfies properties (i) and (iii) of $\mathcal{R}_0$. By the boundedness of b, σ, ℓ, letting ϕ be the identity function, we also have property (ii). Thus, $\Phi(\mu) \in \mathcal{R}_0$. Furthermore, taking a converging sequence $(\mathbb{P}_n)$ in $\mathcal{R}(\mu)$ with limiting measure $\mathbb{P}$, we have $\mathbb{P}$ is also in $\mathcal{R}(\mu)$ since ϕ and $\mathcal{H}_u\phi$ are bounded for each $\phi \in C_c^\infty(\mathbb{R})$ and any $u \in I$. Hence, $\mathcal{R}(\mu)$ is closed for each $\mu \in \mathcal{R}_0$. Therefore, $\mathcal{R}(\mu)$ is nonempty and compact.

Continuity of $\mathcal{R}(\mu)$. We now analyze the continuity of the map $\mu \mapsto \mathcal{R}(\mu)$.

Upper hemicontinuity: Since $\mathcal{R}(\mu)$ is closed for each $\mu \in \mathcal{R}_0$, it suffices to show that its graph is closed. Suppose $\mu_n \to \mu$ in $\mathcal{P}_{\text{Unif}}([0,1] \times \mathcal{D})$ and $\mathbb{P}_n \to \mathbb{P}$ in $\mathcal{P}(\Omega)$. Note that for each $t \in [0, T]$, $\mathcal{H}_u\phi$ is jointly continuous in (x, a). Combining this with the boundedness assumption, by [160, Corollary A.5], we have for each (μ, u), $(m, x) \mapsto M_t^{\mu,\phi}(m, u, x)$ is continuous. Further, since for each (t, x, a), $\mathcal{H}_u\phi(t, x, x', a)$ is continuous in x', and for each $t \in [0, T]$ and a.e. $u \in [0, 1]$, $\mu_t \mapsto \Lambda\mu_t$ is also continuous by [162, Lemma 4.2]. We have for a.e. $u \in [0, 1]$, $(m, x, \mu) \mapsto M_t^{\mu,\phi}(m, u, x)$ is continuous. It follows that for any bounded continuous $\mathcal{F}_s$-measurable function h and all $0 \leqslant s \leqslant t \leqslant T$,

$$\langle \mathbb{P}, (M_t^{\mu,\phi}(m, u, x) - M_s^{\mu,\phi}(m, u, x))h\rangle = \lim_{n\to\infty} \langle \mathbb{P}_n, (M_t^{\mu_n,\phi}(m, u, x) - M_s^{\mu_n,\phi}(m, u, x))h\rangle,$$

which shows that $M_t^{\mu,\phi}$ is a $\mathbb{P}$-martingale, and thus $\mathbb{P} \in \mathcal{R}(\mu)$. The graph of $\mu \mapsto \mathcal{R}(\mu)$ is closed.

Lower hemicontinuity: By [113, Theorem 8.6], we can find a measurable function $\bar{\sigma} : [0, T] \times \mathbb{R} \times \mathcal{P}(\mathbb{R}) \times \mathcal{P}(A) \mapsto \mathbb{R}$ such that $\bar{\sigma}(t, x, \mu, q)$ is continuous for each t,

$$\bar{\sigma}^2(t, x, \mu, q) = \int_A \sigma^2(t, x, \mu, a)q(da),$$

and $\bar{\sigma}(t,x,\mu,\delta_a)=\sigma(t,x,\mu,a)$. Further, we can find a filtered probability space $(\Omega_1,\mathcal{F}_1,\mathbb{F}_1,\mathbb{P}_1)$ that supports a family of independent Brownian motions $W_u, u\in I$, a family of independent random Poisson measures $N_u(dt,de): u\in I$ and a family of $\mathcal{P}(A)$-valued processes $m_u: u\in I$ such that

$$\begin{aligned} dX_u(s) &= \int_A b(s,X_u(s),\Lambda\mu_s(u),a)m_{u,s}(da)ds \\ &\quad + \bar{\sigma}(s,X_u(s),\Lambda\mu_s(u),a)m_{u,s}(da)dW_u(s) \\ &\quad + \int_E\int_A \ell(s,X_u(s),e,a)m_{u,s}(da)\widetilde{N}_u(ds,de), \qquad X_u(0)=\xi_u \qquad u\in I, \end{aligned} \tag{6.12}$$

with $\mathbb{P}=\mathbb{P}_1\circ(dtm_{u,t}(da),du,X_u)^{-1}$. Then on the same probability space, we can write

$$\begin{aligned} dX_u^{(n)}(s) &= \int_A b(s,X_u^{(n)}(s),\Lambda\mu_{n,s}(u),a)m_{u,s}(da)ds \\ &\quad + \bar{\sigma}(s,X_u^{(n)}(s),\Lambda\mu_{n,s}(u),a)m_{u,s}(da)dW_u(s) \\ &\quad + \int_E\int_A \ell(s,X_u^{(n)}(s),e,a)m_{u,s}(da)\widetilde{N}_u(ds,de), \qquad X_u^{(n)}(0)=\xi_u \qquad u\in I, \end{aligned}$$

With μ and μ_n known, the above SDE system has a unique solution thanks to the Lipschitz assumption on b,σ, and ℓ, see e.g., [180]. Then by applying the Burkholder–Davis–Gundy inequality and the boundedness and Lipschitz continuity of b,σ, and ℓ, we have for each $u\in[0,1]$, $\mathbb{E}^{\mathbb{P}_1}[\sup_{t\in[0,T]}|X_u(t)-X_u^{(n)}(t)|^2]\to 0$, which implies that $\mathbb{P}_n\to\mathbb{P}$ in $\mathcal{P}(\Omega)$. We can conclude by using Itô's formula to check that $\mathbb{P}_n\in\mathcal{R}(\mu_n)$. Hence, $\mu\mapsto\mathcal{R}(\mu)$ is lower hemicontinuous.

Analysis of $\mathcal{R}^*(\mu)$. By similar arguments as above, it can be verified that $\Gamma^\mu(m,u,x)$ is continuous in (m,x) for each $u\in I$, due to the continuity of f and g in (x,a). Moreover, for each $t\in[0,T]$, the weak limit μ_t of $\mu_{n,t}$ satisfies that the weak limit of $\Lambda\mu_{n,t}$ is $\Lambda\mu_t$. Hence, the expected functional $(\mu,\mathbb{P})\mapsto\langle\Gamma^\mu,\mathbb{P}\rangle$ is continuous. By applying the famous Berge's Maximum Theorem, the set-valued function $\mathcal{R}^*(\mu)$ is upper hemicontinuous. Then, by the continuity of $\langle\Gamma^\mu,\mathbb{P}\rangle$ and the closedness of $\mathcal{R}(\mu)$, $\mathcal{R}^*(\mu)$ is also closed for each μ. It therefore follows that the graph of $\mathcal{R}^*$ is closed.

Convexity of Φ. Since the map $\mathbb{P}\mapsto\langle\Gamma^\mu,\mathbb{P}\rangle$ is linear in $\mathbb{P}$ and the set $\mathcal{R}(\mu)$ is convex, by the above analysis, we can easily conclude that $\mathcal{R}^*(\mu)$ is convex for each μ. Additionally, using the linearity of $\mathbb{P}\mapsto\mathbb{P}\circ(U,X)^{-1}$, we can also see that $\Phi(\mu)$ is convex.

Now we have verified all the conditions needed to apply Kakutani-Fan-Glicksberg fixed point theorem. Then by the theorem, the set of fixed points of Φ is nonempty and thus there exists at least one graphon equilibrium.

Construction of a strict control

Following the strategy for constructing the Markovian control as in the proof of [162, Theorem 3.2], by using [113, Theorem 8.6], we may find a measurable function $\bar{\sigma}:[0,T]\times\mathbb{R}\times\mathcal{P}(\mathbb{R})\times A\mapsto\mathbb{R}$ such that $\bar{\sigma}(t,x,\mu,q)$ is continuous for each t,

$$\bar{\sigma}^2(t,x,\mu,q)=\int_A \sigma^2(t,x,\mu,a)q(da),$$

and $\bar{\sigma}(t,x,\mu,\delta_a) = \sigma(t,x,\mu,a)$ for each (t,x,μ,a). Further, we can find a filtered probability space $(\Omega_1, \mathcal{F}_1, \mathbb{F}_1, \mathbb{P}_1)$ that supports a family of independent Brownian motions $\{W_{1,u}, u \in I\}$, a family of independent random Poisson measures $\{N_{1,u}(dt,de) : u \in I\}$ and a family of $\mathcal{P}(A)$-valued processes $\{m_{u,t} : u \in I\}$ such that

$$\begin{aligned} dX_{1,u}(s) =& \int_A b(s, X_{1,u}(s), \Lambda\mu_s(u), a) m_{u,s}(da) ds \\ &+ \bar{\sigma}(s, X_{1,u}(s), \Lambda\mu_s(u), m_{u,s}) dW_{1,u}(s) \\ &+ \int_E \int_A \ell(s, X_{1,u}(s), e, a) m_{u,s}(da) \widetilde{N}_{1,u}(ds,de), \qquad X_u(0) = \xi_u \qquad u \in I, \end{aligned} \tag{6.13}$$

and $Q_1 = \mathbb{P}_1 \circ (dt m_{u,t}(da), du, X_{1,u})^{-1}$.

Following the idea in the proof of [160, Theorem 3.7], where the result is shown by using the Mimicking theorem from [71], we can mimic the marginal distribution of the above jump diffusion by using the Markovian projection results for jump diffusion. To do this, we introduce the following combined 2-dimensional process $(U_1(s), X_1(s))_{s\in[0,T]}$ under canonical coupling,

$$\begin{aligned} dX_1(s) =& \int_A b(s, X_1(s), \Lambda\mu_s(U_1(s)), a) m_{U_1(s),s}(da) ds \\ &+ \bar{\sigma}(s, X_1(s), \Lambda\mu_s(U_1(s)), m_{U_1(s),s}) dW_1(s), \\ &+ \int_E \int_A \ell(s, X_1(s), \Phi^{-1}_{U_1(s)}(e-1), a) m_{u,s}(da) \widetilde{\mathcal{N}}_1(ds,de) \\ dU_1(s) =& 0ds + 0dW(s) + \int_E 0\widetilde{\mathcal{N}}_1(ds,de), \end{aligned}$$

Let $\widehat{b}$, $\widehat{\sigma}$ and $\widehat{\ell}$ be defined as

$$\widehat{b}(s, X(s), U(s)) := \mathbb{E}_1\Big[\int_A b(s, X(s), \Lambda\mu_s(U(s)), a) m_{U(s),s}(da) | (U(s), X(s))\Big],$$

$$\widehat{\sigma}^2(s, X(s), U(s)) := \mathbb{E}_1\Big[\int_A \sigma^2(s, X(s), \Lambda\mu_s(U(s)), a) m_{U(s),s}(da) | (U(s), X(s))\Big],$$

$$\widehat{\ell}(s, X(s), U(s), e) := \mathbb{E}_1\Big[\int_A \ell(s, X(s), \Phi^{-1}_{U(s)}(e-1), a) m_{U(s),s}(da) | (U(s), X(s))\Big],$$

and

$$\widehat{f}(s, X(s), U(s)) := \mathbb{E}_1\Big[\int_A f(s, X(s), \Lambda\mu_s(U(s)), a) m_{U(s),s}(da) | (U(s), X(s))\Big].$$

Then, by [59, Theorem 2], there exists another filtered probability space $(\Omega_2, \mathcal{F}_2, \mathbb{F}_2, \mathbb{P}_2)$ that supports a Brownian motion W_2, a Poisson random measure $\mathcal{N}_2(dt,de)$ with the compensator $\widetilde{\mathcal{N}}_2$, such that

$$\begin{aligned} dX_2(s) =& \widehat{b}(s, X_2(s), U_2(s)) ds + \widehat{\sigma}(s, X_2(s), U_2(s)) dW_2(s), \\ &+ \int_E \widehat{\ell}(s, X_2(s), U_2(s), e) \widetilde{\mathcal{N}}_2(ds,de), \\ dU_2(s) =& 0ds + 0dW_2(s) + \int_E 0\widetilde{\mathcal{N}}_2(ds,de), \end{aligned} \tag{6.14}$$

and $(X_2, U_2) \stackrel{d}{=} (X_1, U_1)$.

Recall that according to Assumption 6.1, for all $(t, x, u, \mu) \in [0, T] \times \mathbb{R} \times [0, 1] \times \mathcal{M}^+_{\text{Unif}}([0, 1] \times \mathbb{R})$, there exists some $e^* \in E$ such that the set $K_{e^*}[\mu](t, x, u)$ is convex. This set is also closed since we have assumed that b, σ, ℓ and f are all continuous in a. Note that if $u \mapsto \mu_u$ is measurable, then $u \mapsto \Lambda\mu(u)$ is also measurable, and by definition, we have $(\widehat{b}, \widehat{\sigma}^2, \widehat{\ell}, \widehat{f}) \in K_{e^*}[\mu](t, x, u)$ for each (t, x, u, μ). By applying a measurable selection result [138, Theorem A.9], there exists a measurable function $\alpha^\star : [0, T] \times \mathbb{R} \times [0, 1] \mapsto A$ such that for all $(t, x, u) \in [0, T] \times \mathbb{R} \times [0, 1]$, we have

$$\begin{aligned}
\widehat{b}(t, x, u) =& b(t, x, \Lambda\mu_t(u), a^\star_I(t, x, u)),\\
\widehat{\sigma}^2(t, x, u) =& \sigma^2(t, x, \Lambda\mu_t(u), a^\star_I(t, x, u)),\\
\widehat{\ell}(t, x, u, e^*) =& \ell(t, x, u, e^*, a^\star_I(t, x, u)),
\end{aligned}$$

and

$$\widehat{f}(t, x, u) \leqslant f(t, x, \Lambda\mu_t(u), a^\star_I(t, x, u)).$$

Notice by the second point of Assumption 6.1 and Remark 6.1, $K_e[\mu](t, x, u)$ is convex for all $e \in E$. We can apply [138, Theorem A.9] to the set $K_e[\mu](t, x, u)$ for all $e \in E$. Besides, $\ell(t, x, e, a)$ is affine in e also guarantee for all $e \in E$,

$$\widehat{\ell}(t, x, u, e) = \ell(t, x, u, e, a^\star_I(t, x, u)).$$

We can then rewrite (6.14) as

$$\begin{aligned}
dX_2^{\alpha^\star}(s) =& b(s, X_2^{\alpha^\star}(s), \Lambda\mu_s(U_2(s)), \alpha^\star(s, X_2^{\alpha^\star}(s), U_2(s)))ds\\
&+ \sigma(s, X_2^{\alpha^\star}(s), \Lambda\mu_s(U_2(s)), \alpha^\star(s, X_2^{\alpha^\star}(s), U_2(s)))dW_2(s),\\
&+ \int_E \ell(s, X_2^{\alpha^\star}(s), U_2(s), e, \alpha^\star(s, X_2^{\alpha^\star}(s), U_2(s)))\widetilde{\mathcal{N}}_2(ds, de),\\
dU_2(s) =& 0ds + 0dW_2(s) + \int_E 0\widetilde{\mathcal{N}}_2(ds, de).
\end{aligned} \tag{6.15}$$

Now, $\alpha^\star \in \mathcal{A}_I$, and we have $(X_2^{\alpha^\star}, U_2) \stackrel{d}{=} (X^2, U^2)$ by the weak uniqueness of (6.14).

Define $Q_2 := \mathbb{P}_2 \circ (dt\delta_{\alpha^\star(t, X_2(t), U_2(t))}(da), U_2, X_2)^{-1}$. Then $Q_2 \in \mathcal{R}(\mu)$. Let μ^2 be the joint law of (U_2, X_2), and we have

$$\begin{aligned}
J_G(\mu, \alpha^\star) &= \langle \Gamma^\mu, Q_2 \rangle\\
&= \mathbb{E}_2\Big[\int_0^T f(t, X_2(t), \Lambda\mu_t(U_2(t)), \alpha^\star(t, X_2(t), U_2(t)))dt + g(X_2(T), \Lambda\mu_T(U_2(T))\Big]\\
&= \mathbb{E}_2\Big[\int_0^T f(t, X_2(t), \Lambda\mu^2_t(U_2(t)), \alpha^\star(t, X_2(t), U_2(t)))dt + g(X_2(T), \Lambda\mu^2_T(U_2(T))\Big]\\
&\geqslant \mathbb{E}_2\Big[\int_0^T \widehat{f}(t, X_2(t), U_2(t))dt + g(X_2(T), \Lambda\mu^2_T(U_2(T))\Big]\\
&= \mathbb{E}_1[\int_0^T \int_A f(s, X_1(s), \Lambda\mu_s(U_1(s)), a)m_{U_1(s), s}(da)dt + g(X_1(T), \Lambda\mu_T(U_1(T)))]\\
&= \langle \Gamma^\mu, Q_1 \rangle.
\end{aligned}$$

Notice that for any $Q_1 \in \mathcal{R}^\star(\mu)$, $\langle \Gamma^\mu, Q_1 \rangle \geqslant \langle \Gamma^\mu, Q \rangle$ for any $Q \in \mathcal{R}(\mu)$. This leads to the conclusion

$$J_G(\mu, \alpha^\star) \geqslant \langle \Gamma^\mu, Q_1 \rangle = \sup_{m \in \mathcal{V}} J_G(\mu, m) \geqslant \sup_{\alpha \in \mathcal{A}_I} J_G(\mu, \alpha),$$

with μ chosen to be the solution of the fixed point equation $\Phi(\mu) = \mu$, which means μ is the law of $X^{\alpha^\star}$ under the strict control $\alpha^\star$. $\square$

6.6.2 Stability of graphon: Proof of Theorem 6.8

We will utilize similar techniques for estimates as those in [47] and in Chapter 5. Define $\alpha_u(t) := \alpha(t, u, X_u(t))$ and $\alpha_u^{(n)}(t) := \alpha(t, u, X_u^{(n)}(t))$. By Burkholder–Davis–Gundy inequality, we have

$$\begin{aligned}
&\|X_u^{(n)} - X_u\|_{\mathbb{S}_T^2}^2 \\
&\leqslant C \int_0^T \mathbb{E}\Big|\int_I \int_{\mathbb{R}} G_n(u,v) b(s, X_u^{(n)}(s), x, \alpha_u^{(n)}(s)) \mu_{v,s}^{(n)}(dx) dv \\
&\qquad\qquad - \int_I \int_{\mathbb{R}} G(u,v) b(s, X_u(s), x, \alpha_u(s)) \mu_{v,s}(dx) dv\Big|^2 ds \\
&\quad + C \int_0^T \mathbb{E}\Big|\int_I \int_{\mathbb{R}} G_n(u,v) \sigma(s, X_u^{(n)}(s), x, \alpha_u^{(n)}(s)) \mu_{v,s}^{(n)}(dx) dv \\
&\qquad\qquad - \int_I \int_{\mathbb{R}} G(u,v) \sigma(s, X_u(s), x, \alpha_u(s)) \mu_{v,s}(dx) dv\Big|^2 ds \\
&\quad + C\mathbb{E} \int_0^T \int_E \Big|\ell(s, X_u^{(n)}(s), e, \alpha_u^{(n)}(s)) - \ell(s, X_u(s), e, \alpha_u(s))\Big|^2 N_u(ds, de) + C\mathbb{E}\|\xi_u^{(n)} - \xi_u\|^2.
\end{aligned} \tag{6.16}$$

We calculate the first term; by adding and subtracting terms, we obtain:

$$\begin{aligned}
&\int_0^T \mathbb{E}\Big|\int_I \int_{\mathbb{R}} G_n(u,v) b(s, X_u^{(n)}(s), x, \alpha_u^{(n)}(s)) \mu_{v,s}^{(n)}(dx) dv \\
&\qquad\qquad - \int_I \int_{\mathbb{R}} G(u,v) b(s, X_u(s), x, \alpha_u(s)) \mu_{v,s}(dx) dv\Big|^2 ds \\
&\leqslant C \int_0^T \mathbb{E}\Big[\Big|\int_I (\int_{\mathbb{R}} b(s, X_u(s), x, \alpha_u(s)) \mu_{v,s}(dx))(G(u,v) - G_n(u,v)) dv\Big|^2\Big] ds \qquad (6.17)\\
&\quad + C \int_0^T \mathbb{E}\Big[\int_I \int_{\mathbb{R}} \Big|b(s, X_u(s), x, \alpha_u(s)) - b(s, X_u^{(n)}(s), x, \alpha_u^{(n)}(s))\Big|^2 G_n^2(u,v) \mu_{v,s}(dx) dv\Big] ds \\
&\quad + C \int_0^T \mathbb{E}\Big[\int_I \Big|\int_{\mathbb{R}} b(s, X_u^{(n)}(s), x, \alpha_u^{(n)}(s)) G_n(u,v) [\mu_{v,s} - \mu_{v,s}^{(n)}](dx)\Big|^2 dv\Big] ds.
\end{aligned}$$

Denote the three terms on the right-hand side of inequality (6.17) as $\mathcal{I}_u^{(n),1}$, $\mathcal{I}_u^{(n),2}$, and $\mathcal{I}_u^{(n),3}$ respectively. By utilizing the Lipschitz property of f and α, the property of $X \in \mathcal{M}$, we obtain the following for all $u \in I$:

$$\begin{aligned} C\int_0^T \mathbb{E}\Big[\Big|\int_{\mathbb{R}} b(s, X_u(s), x, \alpha_u(s))\mu_{v,s}(dx)\Big|^2\Big] \\ \leqslant 2C\int_0^T \mathbb{E}\Big[1 + |X_u(s)|^2 + |X_v(s)|^2 + |Z_u(s)|^2\Big] \\ \leqslant C\sup_{u\in I}\mathbb{E}\Big[\sup_{t\in[0,T]}|X_u(t)|^2\Big] \leqslant C. \end{aligned}$$

Hence, by applying Fubini's theorem, we obtain the following for all $u \in I$:

$$\mathcal{I}_u^{(n),1} \leqslant C\int_I \Big(G(u,v) - G_n(u,v)\Big)^2 dv$$

According to the equivalence between the cut norm and the L^1 operator norm of a graphon, we have

$$\int_I \mathcal{I}_u^{(n),1} du \leqslant C\|G - G_n\|_\square.$$

Then, by employing the Lipschitz property of f and α and using inequality (6.2), we obtain

$$\int_I \mathcal{I}_u^{(n),2} du \leqslant C\int_0^T \int_I \mathbb{E}\Big[|X_u(s) - X_u^{(n)}(s)|^2\Big] duds,$$

and

$$\int_I \mathcal{I}_u^{(n),3} du \leqslant C\int_0^T \int_I (\mathcal{W}_2(\mu_{s,v}, \mu_{s,v}^{(n)}))^2 dv \leqslant C\int_0^T \int_I \mathbb{E}\Big[|X_v(s) - X_v^{(n)}(s)|^2\Big] dv.$$

We can address the second term of (6.16) in the same manner. Now, for the third term of (6.16), by utilizing the Lipschitz property of ℓ, we have

$$\mathbb{E}\int_0^T \int_E \Big|\ell(s, X_u^{(n)}(s), e, \alpha_u^{(n)}(s)) - \ell(s, X_u(s), e, \alpha_u(s))\Big|^2 N_u(ds, de) \leqslant C\int_0^T \mathbb{E}|X_u^{(n)}(s) - X_u(s)|^2 ds. \tag{6.18}$$

By combining all the results above and integrating over I, we obtain:

$$\begin{aligned} \int_I \mathbb{E}\Big[\sup_{t\in[0,T]}|X_u^{(n)}(t) - X_u(t)|^2\Big] du \leqslant & C\Big[\int_I \mathbb{E}|\xi_u^{(n)} - \xi_u|^2 du \\ & + \|G - G_n\|_\square + \int_0^T \int_I \mathbb{E}[\sup_{t\in[0,s]}|X_u^{(n)}(t) - X_u(t)|^2] duds\Big]. \end{aligned}$$

Applying Gronwall's Lemma, we conclude that

$$\int_I \mathbb{E}\Big[\sup_{t\in[0,T]}|X_u^{(n)}(t) - X_u(t)|^2\Big] du \leqslant C\Big[\int_I \mathbb{E}|\xi_u^{(n)} - \xi_u|^2 du + \|G - G_n\|_\square\Big].$$

Taking the supremum over I instead of integrating for (6.16), we can first obtain

$$\sup_{u\in I}\mathcal{I}_u^{(n),1}\leqslant C|\int_I\Big(G(u,v)-G_n(u,v)\Big)dv\leqslant C\|G-G_n\|_{\infty\to\infty}.$$

Then, using similar arguments as above, we have

$$\sup_{u\in I}\|X_u^{(n)}(t)-X_u(t)\|^2_{\mathbb{S}^2_T}\leqslant C\Big[\sup_{u\in I}\mathbb{E}|\xi_u^{(n)}-\xi_u|^2+\|G-G_n\|_{\infty\to\infty}\Big],$$

as desired. □

6.6.3 Proof of Lemma 6.9

We proceed with the proof similarly to the continuity arguments presented in [47] and Chapter 5. Here, we employ a different method to couple the system. By coupling X_{u_1} and X_{u_2} with a common Brownian motion W and allowing N_{u_1} and N_{u_2} to jump simultaneously with jump sizes determined by a joint distribution ν_{u_1,u_2}, we have:

$$\begin{aligned}
&\|X_{u_1}-X_{u_2}\|^2_{\mathbb{S}^2_T}\\
&\leqslant C\int_0^T\mathbb{E}\Big|\int_I\int_{\mathbb{R}}G(u_1,v)b(s,X_{u_1}(s),x,\alpha(s,u_1,X_{u_1}(s))\mu_{v,s}(dx)dv\\
&\qquad-\int_I\int_{\mathbb{R}}G(u_2,v)b(s,X_{u_2}(s),x,\alpha(s,u_2,X_{u_2}(s))\mu_{v,s}(dx)dv\Big|^2ds\\
&\quad+C\int_0^T\mathbb{E}\Big|\int_I\int_{\mathbb{R}}G(u_1,v)\sigma(s,X_{u_1}(s),x,\alpha(s,u_1,X_{u_1}(s)))\mu_{v,s}(dx)dv\\
&\qquad-\int_I\int_{\mathbb{R}}G(u_2,v)\sigma(s,X_{u_2}(s),x,\alpha(s,u_2,X_{u_2}(s))\mu_{v,s}(dx)dv\Big|^2ds\\
&\quad+C\mathbb{E}\int_0^T\int_E\Big|\ell(s,X_{u_1}(s),e_1,\alpha(s,u_1,X_{u_1}(s)))-\ell(s,X_{u_2}(s),e_2,\alpha(s,u_2,X_{u_2}(s))\Big|^2N(ds,d(e_1,e_2))\\
&\qquad+C\mathbb{E}|\xi_{u_1}-\xi_{u_2}|^2,
\end{aligned}$$

where $N(ds,d(e_1,e_2))$ is a random Poisson measure with compensator $dt\nu_{u_1,u_2}(d(e_1,e_2))$ and ν_{u_1,u_2} represents the coupled measure of ν_{u_1} and ν_{u_2}. We allow the coupled measure ν_{u_1,u_2} to be coupled in such a way as to achieve the infimum of $\mathbb{E}_{\nu_{u_1,u_2}}|X_1-X_2|^2$ with $\mathcal{L}(x_1)=\nu_{u_1}$ and $\mathcal{L}(X_2)=\nu_{u_2}$. Then, for the first two terms on the right-hand side, we can easily estimate them using a similar approach as in the proof of Theorem 6.8, by employing the Lipschitz continuity of the control $\alpha(t,u,x)$ in (u,x). Denote by $\mathcal{I}$ the sum of the first two terms in the right-hand side of the above equation; we have

$$\mathcal{I}\leqslant C\int_0^T\mathbb{E}|X_{u_1}(s)-X_{u_2}(s)|^2ds+CT\int_I|G(u_1,v)-G(u_2,v)|dv+CT|u_1-u_2|.$$

For the third term, we have

$$\mathbb{E}\int_0^T\int_E\Big|\ell(s,X_{u_1}(s),e_1,\alpha_{u_1}(s))-\ell(s,X_{u_2}(s),e_2,\alpha_{u_2}(s))\Big|^2N(ds,d(e_1,e_2))$$
$$\leqslant C\int_0^T\mathbb{E}|X_{u_1}(s)-X_{u_2}(s)|^2ds+CT(\mathcal{W}_2(\nu_{u_1},\nu_{u_2}))^2+CT|u_1-u_2|.$$

It follows by Gronwall lemma that

$$\mathbb{E}\|X_{u_1}-X_{u_2}\|^2_{\mathbb{S}^2_T}\leqslant C\mathbb{E}|\xi_{u_1}-\xi_{u_2}|^2+CT\int_I|G(u_1,v)-G(u_2,v)|dv+CT(\mathcal{W}_2(\nu_{u_1},\nu_{u_2}))^2.$$

Now, by taking the infimum over random variables ξ_{u_1} and ξ_{u_2} and combining the corresponding assumptions, we can conclude point (i) and (ii) under the respective continuity conditions and Lipschitz conditions. □

6.6.4 Large population convergence: Proof of Theorem 6.10

Denote $\alpha_i^{(n)}(t):=\alpha(t,\frac{i}{n},X_i^{(n)}(t))$, $\alpha_u(t):=\alpha(t,u,X_u(t))$, and $\widetilde{\alpha}_u^{(n)}(t):=\alpha(t,u,\widetilde{X}_u^{(n)}(t))$. First, we estimate the difference between $X^{(n)}$ and $\widetilde{X}^{(n)}$, where $\widetilde{X}^{(n)}$ is the solution of (6.3) with graphon G_n and initial condition $\xi^{(n)}$. By the Burkholder-Davis-Gundy inequality, we have (for some $C>0$):

$$\begin{aligned}
&\|X_i^{(n)}-\widetilde{X}^{(n)}_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T}\\
&\leqslant C\int_0^T\mathbb{E}\Big|\frac{1}{n}\sum_{j=1}^n\zeta_{ij}^{(n)}b(s,X_i^{(n)}(s),X_j^{(n)}(s),\alpha_i^{(n)}(s))\\
&\qquad\qquad-\int_I\int_{\mathbb{R}}G_n(\frac{i}{n},v)b(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),x,\widetilde{\alpha}^n_{\frac{i}{n}}(s))\mu_{v,s}(dx)dv\Big|^2ds\\
&\quad+C\int_0^T\mathbb{E}\Big|\frac{1}{n}\sum_{j=1}^n\zeta_{ij}^{(n)}\sigma(s,X_i^{(n)}(s),X_j^{(n)}(s),\alpha_i^{(n)}(s))\\
&\qquad\qquad-\int_I\int_{\mathbb{R}}G_n(\frac{i}{n},v)\sigma(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),x,\widetilde{\alpha}^n_{\frac{i}{n}}(s))\mu_{v,s}(dx)dv\Big|^2ds\\
&\quad+C\mathbb{E}\int_0^T\int_E\Big|\ell(s,X_i^{(n)}(s),e,\alpha_i^{(n)}(s))\\
&\qquad\qquad-\ell(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),e,\widetilde{\alpha}^n_{\frac{i}{n}}(s))\Big|^2N_{\frac{i}{n}}(ds,de)+C\mathbb{E}|\xi_i^{(n)}-\xi_{\frac{i}{n}}|^2.
\end{aligned}$$

Let's compute the difference between the first term in the right-hand side of the above equation, and similarly for the second term.

$$
\begin{aligned}
&\mathbb{E}\Big|\frac{1}{n}\sum_{j=1}^{n}\zeta_{ij}^{(n)}b(s,X_i^{(n)}(s),X_j^{(n)}(s),\alpha_i^{(n)}(s))\\
&\qquad\qquad -\int_I\int_{\mathbb{R}}G(\frac{i}{n},v)b(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),x,\widetilde{\alpha}^{n}_{\frac{i}{n}}(s))\mu_{v,s}(dx)dv\Big|^2\\
&\leqslant 3\mathbb{E}\Big|\frac{1}{n}\sum_{j=1}^{n}\zeta_{ij}^{(n)}b(s,X_i^{(n)}(s),X_j^{(n)}(s),\alpha_i^{(n)}(s))-\frac{1}{n}\sum_{j=1}^{n}\zeta_{ij}^{(n)}b(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),\widetilde{X}_{\frac{j}{n}}(s),\widetilde{\alpha}^{(n)}_{\frac{i}{n}}(s))\Big|^2\\
&\quad+3\mathbb{E}\Big|\frac{1}{n}\sum_{j=1}^{n}\zeta_{ij}^{(n)}b(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),\widetilde{X}^{(n)}_{\frac{j}{n}}(s),\widetilde{\alpha}^{(n)}_{\frac{i}{n}}(s))\\
&\qquad\qquad -\int_I\int_{\mathbb{R}}G_n(\frac{i}{n},v)b(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),x,\widetilde{\alpha}^{(n)}_{\frac{i}{n}}(s))\widetilde{\mu}^{(n)}_{\frac{[nv]}{n},s}(dx)dv\Big|^2\\
&\quad+3\mathbb{E}\Big|\int_I\int_{\mathbb{R}}G_n(\frac{i}{n},v)b(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),x,\widetilde{\alpha}^{(n)}_{\frac{i}{n}}(s))\widetilde{\mu}^{(n)}_{\frac{[nv]}{n},s}(dx)dv\\
&\qquad\qquad -\int_I\int_{\mathbb{R}}G_n(\frac{i}{n},v)b(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),x,\widetilde{\alpha}^{n}_{\frac{i}{n}}(s))\widetilde{\mu}^{(n)}_{v,s}(dx)dv\Big|^2\\
&=:3(\mathcal{I}_s^{(n),1}+\mathcal{I}_s^{(n),2}+\mathcal{I}_s^{(n),3}),
\end{aligned}
$$

where $\widetilde{X}^{(n)}$ is the solution of (6.3) with graphon G_n, control α, and initial condition ξ, and $\widetilde{\alpha}^{(n)}_u(t) := \alpha(t,u,\widetilde{X}^{(n)}_u(t))$, and $\widetilde{\mu}^{(n)} := \mathcal{L}(\widetilde{X}^{(n)})$. We follow, by using the law of large numbers, similar arguments as in the proof of [47, Lemma 6.1] and combine the Lipschitz property of $\alpha(t,u,x)$ on x. For the first two terms in the right hand side of the above equation, we have

$$
\frac{1}{n}\sum_{i=1}^{n}\mathcal{I}_s^{(n),1}\leqslant C\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\big|X_i^{(n)}(s)-\widetilde{X}_{\frac{i}{n}}(s)\big|^2,
$$

and $\mathcal{I}_s^{(n),2}\leqslant\frac{C}{n}$.

In addition, since α is Lipschitz continuous, we have, by Lemma 6.9, that $\widetilde{\mu}^{(n)}_{v,s}$ is Lipschitz continuous in Wasserstein-2 distance for any $s\in[0,T]$. It follows that

$$
\begin{aligned}
&\int_I\int_{\mathbb{R}}G_n(\frac{i}{n},v)b(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),x,\widetilde{\alpha}^{(n)}_{\frac{i}{n}}(s))\widetilde{\mu}^{(n)}_{\frac{[nv]}{n},s}(dx)dv\\
&\qquad\qquad -\int_I\int_{\mathbb{R}}G_n(\frac{i}{n},v)b(s,\widetilde{X}^{(n)}_{\frac{i}{n}}(s),x,\widetilde{\alpha}^{(n)}_{\frac{i}{n}}(s))\widetilde{\mu}^{(n)}_{v,s}(dx)dv\\
&\leqslant C\int_I\mathcal{W}_2(\widetilde{\mu}^{(n)}_{\frac{[nv]}{n},s},\widetilde{\mu}^{(n)}_{v,s})dv\leqslant\frac{C}{n}.
\end{aligned}
$$

For the jump term, we have similarly as in (6.18)

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\int_0^T\int_E\Big|\ell(s,X_i^{(n)}(s),e,\alpha_i^{(n)}(s))-\ell(s,\widetilde{X}_{\frac{i}{n}}(s),e,\widetilde{\alpha}_{\frac{i}{n}}(s))\Big|^2N_u(ds,de)$$
$$\leqslant C\int_0^T\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|X_i^{(n)}(s)-X_{\frac{i}{n}}(s)|^2ds.$$

Arguing similarly as in the proof of Theorem 6.8, we have

$$\frac{1}{n}\sum_{i=1}^{n}\|X_i^{(n)}-\widetilde{X}_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T}\leqslant\frac{C}{n}. \tag{6.19}$$

Noticing that in fact, $\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|X^{(n)}i(s)-X\frac{i}{n}(s)|^2\leqslant\max_{i\in[n]}\mathbb{E}|X^{(n)}i(s)-X\frac{i}{n}(s)|^2$, by repeating the above analysis and taking the maximum for $i\in[n]$ instead of the sum, we can obtain even

$$\max_{i\in[n]}\|X_i^{(n)}-\widetilde{X}_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T}\leqslant\frac{C}{n}.$$

On the other hand, by Theorem 6.8, we have

$$\int_I\|\widetilde{X}_u-X_u\|^2_{\mathbb{S}^2_T}du\leqslant C\Big[\int_I\mathbb{E}|\xi_u^{(n)}-\xi_u|^2du+\|G-G_n\|_\square\Big]. \tag{6.20}$$

Combining the above two results (6.19) and (6.20), we have

$$\int_I\|X^{(n)}_{\lceil nu\rceil}-X_u\|^2_{\mathbb{S}^2_T}du\leqslant C\Big[\int_I\mathbb{E}|\xi_u^{(n)}-\xi_u|^2du+\|G-G_n\|_\square+\frac{1}{n}\Big].$$

By Lemma 6.9, if G is Lipschitz continuous, $\mathcal{L}(X_u)$ is also Lipschitz continuous in $\mathcal{W}_{2,T}$; if in addition, $\mathcal{L}(\xi_u)$ is Lipschitz continuous in $\mathcal{W}_{2,T}$. We thus have, under continuity conditions,

$$\frac{1}{n}\sum_{i=1}^{n}\|X_i^{(n)}-X_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T}\leqslant C\Big[\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|\xi_i^{(n)}-\xi_{\frac{i}{n}}|^2+\|G-G_n\|_\square\Big]+o(1),$$

and

$$\max_{i\in[n]}\|X_i^{(n)}-X_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T}\leqslant C\Big[\max_{i\in[n]}\mathbb{E}|\xi_i^{(n)}-\xi_{\frac{i}{n}}|^2+\|G-G_n\|_{\infty\to\infty}\Big]+o(1).$$

Further under Lipschitz condition,

$$\frac{1}{n}\sum_{i=1}^{n}\|X_i^{(n)}-X_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T}\leqslant C\Big[\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|\xi_i^{(n)}-\xi_{\frac{i}{n}}|^2+\|G-G_n\|_\square+\frac{1}{n}\Big],$$

and also the maximum type estimate as in Theorem 6.8,

$$\max_{i\in[n]}\|X_i^{(n)}-X_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T}\leqslant C\Big[\max_{i\in[n]}\mathbb{E}|\xi_i^{(n)}-\xi_{\frac{i}{n}}|^2+\|G-G_n\|_{\infty\to\infty}+\frac{1}{n}\Big].$$

Let's define $\widetilde{M}_i^n := \frac{1}{n}\sum_{j=1}^n \zeta_{ij}^{(n)}\delta_{X_{\frac{j}{n}}}$ and $\widetilde{\Lambda\mu}_i^n := \frac{1}{n}\sum_{j=1}^n G(\frac{i}{n},\frac{j}{n})\mu_{\frac{j}{n}}$. Now suppose G and $\mathcal{L}(\xi_u)$ are Lipschitz continuous. Let $\zeta_{ij}^{(n)}$ satisfy the Regularity assumption for graphon G. For any bounded Lipschitz continuous function h on $\mathcal{D}$, we have again by Lemma 6.9 (ii),

$$\langle h, \widetilde{\Lambda\mu}_i^n\rangle - \langle h, \Lambda\mu(\frac{i}{n})\rangle = \frac{1}{n}\sum_{j=1}^n G(\frac{i}{n},\frac{j}{n})\langle h, \mu_{\frac{j}{n}}\rangle - \int_I G(\frac{i}{n},v)\langle h,\mu_v\rangle dv \leqslant \frac{C}{n}.$$

Moreover, we have

$$\begin{aligned}
&\mathbb{E}\Big[\langle h, \widetilde{M}_i^n\rangle - \langle h, \widetilde{\Lambda\mu}_i^n\rangle\Big]^2 \\
&= \frac{1}{n^2}\mathbb{E}\Big[\sum_{j=1}^n\Big(\zeta_{ij}^{(n)}h(X_{\frac{j}{n}}) - G(\frac{i}{n},\frac{j}{n})\mathbb{E}h(X_{\frac{j}{n}})\Big)\Big]^2 \\
&= \frac{1}{n^2}\sum_{j=1}^n\mathbb{E}\Big[\zeta_{ij}^{(n)}h(X_{\frac{j}{n}}) - G(\frac{i}{n},\frac{j}{n})\mathbb{E}h(X_{\frac{j}{n}})\Big]^2 \\
&\quad + \frac{1}{n^2}\sum_{j=1}^n\sum_{k\neq j}\mathbb{E}\Big[\Big(\zeta_{ij}^{(n)}h(X_{\frac{j}{n}}) - G(\frac{i}{n},\frac{j}{n})\mathbb{E}h(X_{\frac{j}{n}})\Big)\Big(\zeta_{ik}^{(n)}h(X_{\frac{k}{n}}) - G(\frac{i}{n},\frac{k}{n})\mathbb{E}h(X_{\frac{k}{n}})\Big)\Big] \\
&\leqslant \frac{\|h\|_\infty}{n},
\end{aligned}$$

where the last inequality comes from the boundedness of h and the independence of $\zeta_{ij}^{(n)}$. On the other hand, by previous result, we have

$$\mathbb{E}\Big[\langle h, M_i^{(n)}\rangle - \langle h, \widetilde{M}_i^n\rangle\Big]^2 \leqslant \frac{C}{n}\sum_{j=1}^n \|X_j^{(n)} - X_{\frac{j}{n}}\|^2_{\mathbb{S}^2_T}.$$

Combine the above three results, we can conclude that for any $i\in[n]$,

$$\mathbb{E}\Big[\langle h, M_i^{(n)}\rangle - \langle h, \Lambda\mu(\frac{i}{n})\rangle\Big]^2 \leqslant \frac{C}{n}\sum_{j=1}^n \|X_j^{(n)} - X_{\frac{j}{n}}\|^2_{\mathbb{S}^2_T} + \frac{C}{n}.$$

Notice that in the assumptions of the theorem, $\zeta^{(n)}$ satisfies regularity Assumption 6.5 with intermediate step graphon G_n, not for G. Recall the definition of $\widetilde{\mu}^{(n)}$ which was defined before. We have for any $i\in[n]$,

$$\mathbb{E}\Big[\langle h, M_i^{(n)}\rangle - \langle h, \Lambda\widetilde{\mu}^{(n)}(\frac{i}{n})\rangle\Big]^2 \leqslant \frac{C}{n}\sum_{j=1}^n \|X_j^{(n)} - X_{\frac{j}{n}}\|^2_{\mathbb{S}^2_T} + \frac{C}{n}.$$

Finally, by the stability of graphon in Theorem 6.8, we have

$$\begin{aligned}
&\langle h, \Lambda\widetilde{\mu}^{(n)}(\frac{i}{n})\rangle - \langle h, \Lambda\mu(\frac{i}{n})\rangle \\
&\qquad\leqslant \int_I G_n(\frac{i}{n},v)\langle h, \widetilde{\mu}_v^{(n)}\rangle dv - \int_I G(\frac{i}{n},v)\langle h,\mu_v\rangle dv \\
&\qquad\leqslant C\int_I\Big(G_n(\frac{i}{n},v) - G(\frac{i}{n},v)\Big)dv + C\|G_n - G\|_\square,
\end{aligned}$$

Hence, we finally arrive at the conclusion that for any $i \in [n]$,

$$\mathbb{E}\Big[\langle h, M_i^{(n)}\rangle - \langle h, \Lambda\mu(\frac{i}{n})\rangle\Big]^2 \leqslant \frac{C}{n}\sum_{j=1}^{n}\mathbb{E}|\xi_j^{(n)} - \xi_{\frac{j}{n}}|^2 + C\|G_n - G\|_{\Box} + C\int_I \Big(G_n(\frac{i}{n}, v) - G(\frac{i}{n}, v)\Big)dv + \frac{C}{n}. \tag{6.21}$$

If $|G_n - G|_{\infty\to\infty} = O(n^{-1})$ and $\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}|\xi_j^{(n)} - \xi_{\frac{j}{n}}|^2 = O(n^{-1})$, by Markov's inequality, we can conclude that for all $i \in [n]$, as $n \to \infty$, $M_i^{(n)} \to \Lambda\mu(\frac{i}{n})$ in probability in the weak sense. Moreover, we have

$$\langle h, M_i^{(n)}\rangle - \langle h, \Lambda\mu(\frac{i}{n})\rangle = O_p(n^{-\frac{1}{2}}).$$

Finally, by the Lipschitz continuity of the law of X_u in u (Lemma 6.9 (ii)) and the Lipschitz continuity of graphon G, we have that $u \mapsto \Lambda\mu(u)$ is also Lipschitz continuous. Hence, we have

$$\frac{1}{n}\sum_{i=1}^{n}\langle h, M_i^{(n)}\rangle - \int_I \langle h, \Lambda\mu(v)\rangle dv = O_p(n^{-\frac{1}{2}}).$$

The corresponding results under continuity assumptions follow from the above analysis. Thus, the proof is complete. □

6.6.5 Convergence for general graphon: Proof of Corollary 6.13

We only highlight here the necessary changes from the continuous case (Theorem 6.10) to the general case. We will keep the same notation as in the proof of Theorem 6.10. By our assumptions, $(\frac{i}{n}, \frac{j}{n})_{i,j\in[n]}$ is in the continuous modification set of G. By Lusin's Theorem, we can approximate G by a continuous graphon $\bar{G}$ such that $\bar{G}(\frac{i}{n}, \frac{j}{n}) = G(\frac{i}{n}, \frac{j}{n})$ and $|\bar{G} - G|_{L^1} = 0$. Let $\bar{X}$ be the solution of (6.3) associated with control α and graphon $\bar{G}$. By Theorem 6.8, $\int_I \mathcal{W}_{2,T}(\mu_u, \bar{\mu}_u)du = 0$ and by Lemma 6.9, $\bar{\mu}_u := \mathcal{L}(\bar{X}_u)$ is continuous in Wasserstein-2 distance. Thus we have

$$\frac{1}{n}\sum_{i=1}^{n}\|\bar{X}_{\frac{i}{n}} - X_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T} = \int_I \|\bar{X}_{\frac{[nu]}{n}} - X_{\frac{[nu]}{n}}\|^2_{\mathbb{S}^2_T} du \leqslant \int_I \|\bar{X}_u - X_u\|^2_{\mathbb{S}^2_T} du + o(1).$$

Notice that the last inequality comes from the assumption that $(\frac{i}{n}, \frac{j}{n})_{i,j\in[n]}$ is in the continuous modification set of G, which means that $\mu_{\frac{i}{n}}$ is continuous on a subset of I with Lebesgue measure 1. On the other hand, by the results of Theorem 6.10, we have

$$\frac{1}{n}\sum_{i=1}^{n}\|X_i^{(n)} - \bar{X}_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T} \leqslant C\Big[\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_{\Box}\Big] + o(1).$$

Combining the above two formulas, we obtain

$$\frac{1}{n}\sum_{i=1}^{n}\|X_i^{(n)} - X_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T} \leqslant C\Big[\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_{\Box}\Big] + o(1).$$

Furthermore, using a similar idea to the graphon G, we can extend the convergence of the empirical neighborhood measure for non-continuous graphons G by arguing similarly as in the proof of [47, Theorem 3.1] by using a sequence of step graphons G_m to approximate G. Indeed, by the assumption that $(\frac{i}{n}, \frac{j}{n})_{i,j\in[n]}$ is in the continuous modification set of G, we can find a sequence of step graphons G_m such that

$$\|G_m - G\|_\square \leqslant C(m),$$

and for all $i \in [n]$,

$$\int_I \Big(G_m(\frac{i}{n}, v) - G(\frac{i}{n}, v)\Big) dv \leqslant C_1(m).$$

Denote by μ^m the law of solution of (6.3) with graphon G_m and the same initial condition and control as X. Then all the analysis above still applies with each G_m, and we similarly have

$$\langle h, \Lambda\mu^m(u)\rangle - \langle h, \Lambda\mu(u)\rangle \leqslant C\int_I \Big(G_m(\frac{i}{n}, v) - G(\frac{i}{n}, v)\Big) dv + C\|G_m - G\|_\square.$$

Combine (6.21) for graphon G_m, we have for each $i \in [n]$,

$$\begin{aligned}\mathbb{E}\Big[\langle h, M_i^{(n)}\rangle - \langle h, \Lambda\mu(\frac{i}{n})\rangle\Big]^2 \leqslant \frac{C}{n}\sum_{j=1}^n \mathbb{E}|\xi_j^{(n)} - \xi_{\frac{j}{n}}|^2 + C\|G_n - G_m\|_\square + C\|G_n - G_m\|_{\infty\to\infty} \\ + \frac{C}{n} + C\int_I \Big(G_m(\frac{i}{n}, v) - G(\frac{i}{n}, v)\Big) dv + C\|G_m - G\|_\square.\end{aligned} \tag{6.22}$$

Hence we can conclude by first letting n tend to ∞, and then letting m tend to ∞. □

6.6.6 Stability of control: Proof of Lemma 6.17

We will argue via Jensen's inequality and the Markovian projection for our controlled jump diffusion. Recall the definition of $\mathcal{R}(\mu)$ in 6.21 and denote its u-marginal by $\mathcal{R}_u(\mu)$. We first prove that an optimizer for

$$\big\{\langle \Gamma^\mu, \mathbb{P}\rangle : \mathbb{P} \in \mathcal{R}_u(\mu)\big\}$$

must be of the form $\mathbb{P} = \mathcal{L}(dt\delta_{\alpha(t,X_t^\alpha)}(da), \delta_u, \delta_{X^\alpha})$. For two different functions $\alpha_1 : [0,T] \times \mathbb{R} \to A$ and $\alpha_2 : [0,T] \times \mathbb{R} \to A$, and arbitrary $\lambda \in (0,1)$, define the λ-averaged control as

$$\bar{\alpha}_\lambda(t,x) := \lambda\alpha_1(t,x) + (1-\lambda)\alpha_2(t,x).$$

Denote for short by X the state dynamic governed by the relaxed control $\bar{\delta}_\lambda$ defined by

$$\bar{\delta}_\lambda(t,X) = \lambda\delta_{\alpha_1(t,X)} + (1-\lambda)\delta_{\alpha_2(t,X)}.$$

Define also $\bar{\mathbb{P}}_\lambda := \mathcal{L}(dt\delta_{\bar{\alpha}_\lambda(t,X^{\bar{\alpha}_\lambda})}(da), \delta_u, \delta_{X^{\bar{\alpha}_\lambda}})$ and $\widetilde{\mathbb{P}}_\lambda := \mathcal{L}(dt\bar{\delta}_\lambda(t,X)(da), \delta_u, \delta_X)$. It then suffices to prove

$$\langle \Gamma^\mu, \bar{\mathbb{P}}_\lambda\rangle > \langle \Gamma^\mu, \widetilde{\mathbb{P}}_\lambda\rangle.$$

Let $\overline{X}$ be the Markovian projection process of X which has the same law as X. Notice that, by the analysis in how we construct a strict control under our standing assumption, there exists such a strict control $\widehat{\alpha}_\lambda$ such that $\overline{X}$ is the driven state process under $\widehat{\alpha}_\lambda$, which satisfies

$$\begin{aligned}b(t, X(t), \Lambda\mu_t(u), \widehat{\alpha}_\lambda(t, X(t))) =& \lambda b(t, X(t), \Lambda\mu_t(u), \alpha_1(t, X(t)))\\ &+ (1-\lambda) b(t, X(t), \Lambda\mu_t(u), \alpha_2(t, X(t))),\\ \sigma(t, X(t), \Lambda\mu_t(u), \widehat{\alpha}_\lambda(t, X(t))) =& \lambda \sigma(t, X(t), \Lambda\mu_t(u), \alpha_1(t, X(t)))\\ &+ (1-\lambda)\sigma(t, X(t), \Lambda\mu_t(u), \alpha_2(t, X(t))),\end{aligned}$$

and for any $e \in E$,

$$\begin{aligned}\ell(t, X(t), e, \widehat{\alpha}_\lambda(t, X(t))) =& \lambda \ell(t, X(t), e, \alpha_1(t, X(t)))\\ &+ (1-\lambda)\ell(t, X(t), e, \alpha_2(t, X(t))),\end{aligned}$$

From concavity Assumption 6.6, we have

$$f(t, X(t), \Lambda\mu_t(u), \widehat{\alpha}_\lambda(t, X(t))) \geqslant f(t, X(t), \Lambda\mu_t(u), \bar{\alpha}_\lambda(t, X(t))).$$

By Jensen's inequality, we have

$$\begin{aligned}\langle \Gamma^\mu, \widetilde{\mathbb{P}}_\lambda \rangle = & \mathbb{E}\Big[\int_0^T \lambda f(t, X(t), \Lambda\mu_t(u), \alpha_1(t, X(t))) + (1-\lambda) f(t, X(t), \Lambda\mu_t(u), \alpha_2(t, X(t)))\\ &\qquad + g(X(T), \Lambda\mu_T(u))\Big]\\ &< \mathbb{E}\Big[\int_0^T f(t, X(t), \Lambda\mu_t(u), \bar{\alpha}_\lambda(t, X(t))) + g(X_T, \Lambda\mu_T(u))\Big]\\ &\leqslant \mathbb{E}\Big[\int_0^T f(t, X(t), \Lambda\mu_t(u), \widehat{\alpha}_\lambda(t, X(t))) + g(X_T, \Lambda\mu_T(u))\Big]\\ &= \mathbb{E}\Big[\int_0^T f(t, \overline{X}(t), \Lambda\mu_t(u), \widehat{\alpha}_\lambda(t, \overline{X}(t))) + g(\overline{X}_T, \Lambda\mu_T(u))\Big]\\ &= J_G^{u,\xi_u}(\mu, \widehat{\alpha}_\lambda).\end{aligned}$$

It is also easy to see that the optimizer is unique. Suppose we have two different optimizer controls, α_1 and α_2. By our assumption and using Jensen's inequality, we can find a better control, $\alpha^\star$, which is distinct from both α_1 and α_2, by arguing in a similar manner as above.

Now, denote $\alpha_u^\star$ as the unique optimal control corresponding to $J_G^{u,\xi_u}(\mu, \widehat{\alpha}_\lambda)$ with a given $\mu^\star \in \mathcal{P}_{\text{Unif}}([0,1] \times \mathcal{D})$. Following similar arguments as in the proof of [162, Lemma 5.2], we can establish the continuity of $\mathcal{L}(X^{\alpha_u^\star, \xi_u})$ with respect to u. Due to the assumed continuity of G, $\Lambda\mu_t^\star(u)$ is also continuous in u for any $t \in [0,T]$ in the weak sense. Define

$$V_G^u(\mu, \alpha) := \mathbb{E}\Big[\int_0^T f(t, X(t), \Lambda\mu_t(u), \alpha(t, X(t)))dt + g(X(T), \Lambda\mu_T(u))\Big) du\Big],$$

with X admitting the law $\mu_u^\star$. Recall that $\alpha^\star(t,u,x) := \alpha_u^\star(t,x)$. We now prove that $\alpha^\star(t,u,x)$ is continuous in u. Indeed, observing the continuity of $\mu_{u,t}^\star$ and $\Lambda\mu_t^\star(u)$ in u for any $t \in [0,T]$, and by the uniqueness of $\alpha_u^\star$, for arbitrarily small ε_1 and ε_2, there exist $\delta_1(\varepsilon_1) > 0$, $\delta_2(\varepsilon_1) > 0$ which tend to 0 as ε_1 tends to 0, and $\delta_3(\varepsilon_2) > 0$ which tends to 0 as ε_2 tends to 0, such that for $u_1 \in (u-\varepsilon_1, u+\varepsilon_1)$ and $\alpha \notin (\alpha_u^\star - \varepsilon_2, \alpha_u^\star + \varepsilon_2)$, we have

$$\begin{aligned} V_G^{u_1}(\mu^\star,\alpha) \leqslant & V_G^u(\mu^\star,\alpha_u^\star) + \delta_1(\varepsilon_1) - \delta_3(\varepsilon_2), \\ V_G^u(\mu^\star,\alpha_u^\star) \leqslant & V_G^u(\mu^\star,\alpha_{u_1}^\star) + \delta_2(\varepsilon_1), \quad \text{and}, \\ V_G^{u_1}(\mu^\star,\alpha_{u_1}^\star) \geqslant & V_G^u(\mu^\star,\alpha_{u_1}^\star) - \delta_1(\varepsilon_1). \end{aligned} \tag{6.23}$$

From the first two inequalities mentioned above, we can obtain

$$V_G^{u_1}(\mu^\star,\alpha) \leqslant V_G^u(\mu^\star,\alpha_{u_1}^\star) + \delta_2(\varepsilon_1) + \delta_1(\varepsilon_1) - \delta_3(\varepsilon_2). \tag{6.24}$$

Comparing (6.23) and (6.24) and letting $\varepsilon_1 \to 0$, we must have $\alpha_{u_1}^\star \in (\alpha_u^\star - \varepsilon_2, \alpha_u^\star + \varepsilon_2)$. Then, as $\varepsilon_2 \to 0$, we obtain $\alpha_{u_1}^\star(t,x) \to \alpha_u^\star(t,x)$ a.e. for (t,x).

To prove the continuity of $x \mapsto \alpha^\star(t,u,x)$ for each $u \in I$, we proceed by contradiction. If $\alpha^\star(t,u,x)$ is not continuous in x, then the coefficients in the dynamics of $X_t^{\alpha^\star}$ are not continuous, it then suffices to prove that $\mathcal{L}(X^{\alpha_u^\star,\xi_u})$ is not continuous in u. Suppose x^* is a discontinuous point of $\alpha^\star$. If $X_{u_1}^{\alpha_{u_1}^\star}(t) = x^*$ for some $t \in [0,T]$, then for some small $\epsilon > 0$ and $X_{u_2}^{\alpha_{u_2}^\star}(t) \in (x^*-\epsilon, x^*+\epsilon)$ but not equal to x^*, we must have, in a small time interval $[t, t+\Delta t]$, for some $\eta > 0$ not depending on ϵ such that $|\alpha^\star(s,u_1,X_{u_1}^{\alpha_{u_1}^\star}(s)) - \alpha^\star(s,u_2,X_{u_2}^{\alpha_{u_2}^\star}(s))| > \eta$,

$$\begin{aligned} &\int_t^{t+\Delta t}\int_I\int_{\mathbb{R}} G(u_1,v)b(s,X_{u_1}^{\alpha_{u_1}^\star}(s),x,\alpha^\star(s,u_1,X_{u_1}^{\alpha_{u_1}^\star}(s)))\mu_{v,s}^{\alpha_{u_1}^\star}(dx)dvds \\ &\qquad - \int_t^{t+\Delta t}\int_I\int_{\mathbb{R}} G(u_2,v)b(s,X_{u_2}^{\alpha_{u_2}^\star}(s),x,\alpha^\star(s,u_2,X_{u_2}^{\alpha_{u_1}^\star}(s)))\mu_{v,s}^{\alpha_{u_2}^\star}(dx)dvds \\ &\geqslant C\eta. \end{aligned}$$

This shows that at each discontinuous point, the drift amplifies a significant difference. If we take the diffusion and jumps into consideration, the difference will become even larger. By Burkholder-Davis-Gundy inequality, we have $\|X_{u_1}^{\alpha_{u_1}^\star} - X_{u_2}^{\alpha_{u_2}^\star}\|_{\mathbb{S}_T^2} > \eta'$ for some constant η' not depending on u_1, u_2. Since the support of Brownian motion is on the whole $\mathbb{R}$, we have a set $\mathcal{O} \subset \Omega$ with $\mathbb{P}(\mathcal{O}) > 0$ such that for some $u \in I$, $X_u^{\alpha_u^\star}(\omega)$ passes through the discontinuous points of $\alpha^\star$. Then by Skorokhod Representation Theorem, it follows that $\mathcal{L}(X^{\alpha_u^\star,\xi_u})$ cannot be continuous in u.

Using the same arguments and combining the Lipschitz continuity of $f(t,x,\mu,a)$ in (x,μ,a), the Lipschitz continuity of $g(x,\mu)$ in (x,μ), and the Lipschitz continuity of $\mathcal{L}(X_u^\alpha)$ and $\Lambda\mu(u)$ with respect to u, we can conclude the Lipschitz continuity under the corresponding assumptions. $\square$

6.6.7 Proof of Theorem 6.18 and Theorem 6.20

We consider the following decomposition:

$$\epsilon_i^{(n)}(\boldsymbol{u}^{(n)}) \leqslant \sup_{\alpha\in\mathcal{A}_n} \Delta_i^{(n),1}(\alpha,\boldsymbol{u}^{(n)}) + \sup_{\alpha\in\mathcal{A}_n} \Delta_i^{(n),2}(\alpha,\boldsymbol{u}^{(n)}) + \Delta_i^{(n),3}(\boldsymbol{u}^{(n)}),$$

where the three Δ_is are defined as follows:

$$\begin{aligned}\Delta_i^{(n),1}(\alpha,\boldsymbol{u}^{(n)}) :=\ & \mathbb{E}\Big[\int_0^T f(t,X_i^{(n),\alpha,-i}(t),M_i^{(n),-i}(t),\alpha(t))dt + g(X_i^{(n),\alpha,-i}(T),M_i^{(n),-i}(T))\Big]\\ & -\mathbb{E}\Big[\int_0^T f(t,X^{\star,\alpha,-\frac{i}{n}}_{u_i^{(n)}}(t),\Lambda\mu_t^{\alpha,-\frac{i}{n}}(u_i^{(n)}),\alpha(t))dt + g(X^{\star,\alpha,-\frac{i}{n}}_{u_i^{(n)}}(T),\Lambda\mu_T^{\alpha,-\frac{i}{n}}(u_i^{(n)}))\Big],\end{aligned}$$

$$\begin{aligned}\Delta_i^{(n),2}(\alpha,\boldsymbol{u}^{(n)}) :=& \mathbb{E}\Big[\int_0^T f(t,X^{\star,\alpha,-\frac{i}{n}}_{u_i^{(n)}}(t),\Lambda\mu_t^{\alpha,-\frac{i}{n}}(u_i^{(n)}),\alpha(t))dt + g(X^{\star,\alpha,-\frac{i}{n}}_{u_i^{(n)}}(T),\Lambda\mu_T^{\alpha,-\frac{i}{n}}(u_i^{(n)}))\Big]\\ & - \mathbb{E}\Big[\int_0^T f(t,X^{\star}_{u_i^{(n)}}(t),\Lambda\mu_t(u_i^{(n)}),\alpha^\star_{\frac{i}{n}}(t))dt + g(X^\star_{u_i^{(n)}}(T),\Lambda\mu_T(u_i^{(n)}))\Big],\end{aligned}$$

and

$$\begin{aligned}\Delta_i^{(n),3}(\alpha,\boldsymbol{u}^{(n)}) :=& \mathbb{E}\Big[\int_0^T f(t,X^{\star}_{u_i^{(n)}}(t),\Lambda\mu_t(u_i^{(n)}),\alpha^\star_{\frac{i}{n}}(t))dt + g(X^\star_{u_i^{(n)}}(T),\Lambda\mu_T(u_i^{(n)}))\Big]\\ & - \mathbb{E}\Big[\int_0^T f(t,X_i^{(n),\star}(t),M_i^{(n)}(t),\alpha_i^\star(t))dt + g(X_i^{(n),\star}(T),M_i^{(n)}(T))\Big].\end{aligned}$$

Here, $X^{(n),\alpha,-i}$ denotes the state vector of the n-player interacting system with the i-th player choosing control α and the other players keeping the control $\alpha^\star$; $X^{\star,\alpha,-\frac{i}{n}}$ is the state family of the limit graphon system with the label $\frac{i}{n}$ choosing control α and the others keeping the control $\alpha^\star$; $M^{(n),-i}$ is the empirical neighborhood measure induced by $X^{(n),\alpha,-i}$; and $\Lambda\mu^{\alpha,-\frac{i}{n}}$ is the graphon measure induced by $X^{\star,\alpha,-\frac{i}{n}}$.

Lipschitz graphon We concentrate on proving the theorems for Lipschitz continuous graphon. Then, by following similar arguments, the results for continuous graphon are straightforward.

For the Lipschitz graphon, under the conditions assumed in Theorem 6.18, by Lemma 6.17, we first have the equilibrium control $\alpha^\star(t,u,x)$ that is Lipschitz continuous in (u,x). Then, by choosing $u_i^{(n)} = \frac{i}{n}$, we have the following estimate for the difference between the empirical neighborhood measure and the graphon neighborhood measure in the weak convergence sense, for $t\in[0,T]$ and any bounded Lipschitz continuous function h on $\mathbb{R}$,

$$\langle h, M_i^{(n)}(t)\rangle - \langle h,\Lambda\mu_t(\frac{i}{n})\rangle = O_p(n^{-\frac{1}{2}}).$$

Furthermore, we can assume that W_i, N_i, and $W_{\frac{i}{n}}, N_{\frac{i}{n}}$ are the same for each $i \in [n]$, since such a correspondence does not change the law of both systems and will not influence our approximation. We have

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\|X_i^{(n),\star} - X^{\star}_{u_i^{(n)}}\|^2_{\mathbb{S}^2_T} \leqslant C\Big[\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_{\square} + \frac{1}{n}\Big].$$

In the rest of the proof, we will always establish such a correspondence between the finite system and the graphon system. Since $f(t,x,\mu,a)$ and $g(x,\mu)$ are Lipschitz continuous in x and μ, we have

$$\frac{1}{n}\sum_{i=1}^{n}\Delta_i^{(n),3}(\alpha, \boldsymbol{u}^{(n)}) \leqslant C\Big[\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_{\square} + \frac{1}{n}\Big].$$

For the first term $\Delta_i^{(n),1}(\alpha, \boldsymbol{u}^{(n)})$, notice that when n is large, the change in the dynamics of player i will not have a significant impact on the distributions of trajectories of other players, which is also true for the limit graphon system. Hence, for $j \neq i$, one can easily show that

$$\|X_j^{(n),\alpha,-i} - X_j^{(n),\star}\|_{\mathbb{S}^2_T} \leqslant \frac{C}{n},$$

and moreover, for $u \neq \frac{i}{n}$, a.s.

$$\widetilde{X}_u^{\star,\alpha,-\frac{i}{n}} = \widetilde{X}_u^{\star}.$$

Combining the previous result, we have for $j \neq i$,

$$\frac{1}{n}\sum_{j\neq i}\|X_j^{(n),\alpha,-i} - \widetilde{X}^{\star,\alpha,-\frac{i}{n}}_{\frac{j}{n}}\|^2_{\mathbb{S}^2_T} \leqslant C\Big[\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \frac{1}{n}\Big],$$

where $\widetilde{X}^{\star,\alpha,-\frac{i}{n}}$ is the state family of the limit graphon system induced by the step graphon G_n, with label i/n choosing control α and the others keeping the control $\alpha^{\star}$.

Now we will study the behavior of player i under control α. By similar arguments as in the proof of Lemma 6.17 and some calculations, we can show that the following inequality holds:

$$\|X_i^{(n),\alpha,-i} - \widetilde{X}^{\star,\alpha,-\frac{i}{n}}_{\frac{i}{n}}\|^2_{\mathbb{S}^2_T} \leqslant \frac{C}{n}\sum_{j\neq i}\mathbb{E}\|X_j^{(n),\alpha,-i} - X^{\star,\alpha,-\frac{i}{n}}_{\frac{j}{n}}\|^2_{\mathbb{S}^2_T} + \frac{C}{n}.$$

By using the stability of graphon Theorem 6.8 again, we have

$$\int_I \|\widetilde{X}_u^{\star,\alpha,-\frac{i}{n}} - X_u^{\star,\alpha,-\frac{i}{n}}\|^2_{\mathbb{S}^2_T}du \leqslant C\Big[\int_I \mathbb{E}|\xi_u^{(n)} - \xi_u|^2 du + \|G - G_n\|_{\square}\Big].$$

Moreover, we have for all $u \in I$,

$$\|\widetilde{X}_u^{\star,\alpha,-\frac{i}{n}} - X_u^{\star,\alpha,-\frac{i}{n}}\|^2_{\mathbb{S}^2_T} \leqslant C\Big[\mathbb{E}|\xi_u^{(n)} - \xi_u|^2 + |\int_I\Big(G(u,v) - G_n(u,v)\Big)dv|\Big].$$

Finally, by combining the above three results, we obtain:

$$\|X_i^{(n),\alpha,-i} - X_{\frac{i}{n}}^{\star,\alpha,-\frac{i}{n}}\|_{\mathbb{S}_T^2}^2 \leqslant C\Big[\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \frac{1}{n} + \mathbb{E}|\xi_u^{(n)} - \xi_u|^2 + |\int_I \Big(G(u,v) - G_n(u,v)\Big)dv|\Big].$$

Hence, we can similarly obtain

$$\frac{1}{n}\sum_{i=1}^n \Delta_i^{(n),1}(\alpha, \boldsymbol{u}^{(n)}) \leqslant C\Big[\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_\Box + \frac{1}{n}\Big],$$

and further for each $i \in [n]$,

$$\Delta_i^{(n),1}(\alpha, \boldsymbol{u}^{(n)}) \leqslant C\Big[\mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + |\int_I \Big(G(\frac{i}{n}, v) - G_n(\frac{i}{n}, v)\Big)dv| + \frac{1}{n}\Big].$$

Therefore,

$$\max_{i\in[n]} \Delta_i^{(n),1}(\alpha, \boldsymbol{u}^{(n)}) \leqslant C\Big[\max_{i\in[n]} \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_\infty + \frac{1}{n}\Big].$$

For $\Delta_i^{(n),2}(\alpha, \boldsymbol{u}^{(n)})$, it is apparently non positive by the marginal supreme Proposition 6.7. Combining the above three estimates, we can conclude that

$$\frac{1}{n}\sum_{i=1}^n \epsilon_i^{(n)} \leqslant C\Big[\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n\|_\Box + \frac{1}{n}\Big],$$

and for the maximum

$$\max_{i\in[n]} \epsilon_i^{(n)} \leqslant C\Big[\max_{i\in[n]} \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|\int_I \Big(G(\frac{i}{n}, v) - G_n(\frac{i}{n}, v)\Big)dv\|_\infty + \frac{1}{n}\Big].$$

We have proven the results for $u_i^{(n)} = \frac{i}{n}$. It is then easy to generalize the above estimates for all choices $\boldsymbol{u}^{(n)}$ in $\mathcal{I}_1^{(n)} \times \cdots \times \mathcal{I}_n^{(n)}$. Since $\alpha^\star(t,u,x)$ is Lipschitz in u, we have for some constant C and any $u_i^{(n)} \in \mathcal{I}_i^{(n)}$, for each i,

$$\alpha^\star(t, u_i^{(n)}, x) = \alpha^\star(t, \frac{i}{n}, x) + C\frac{1}{n}.$$

Thus following similar arguments as in the proof of Theorem 6.10, we can recover the results in Theorem 6.10 with control $\alpha_i^{(n)}(t,x) = \alpha^\star(t, u_i^{(n)}, x)$. Hence all analysis above can apply. The proof is complete now.

Sampling graphon We now proceed with the proof of Theorem 6.20.

Let $G_n(\boldsymbol{U}^{(n)})$ be the step graphon generated by the sampling graphon G, i.e

$$G_n(\boldsymbol{U}^{(n)})(u,v) = G\big(U^{(n)}_{(\lceil nu\rceil)}, U^{(n)}_{(\lceil nv\rceil)}\big).$$

Let $\widetilde{X}^{(n)}(\boldsymbol{U}^{(n)})$ be the solution of (6.3) with graphon $G_n(\boldsymbol{U}^{(n)})$, control α and initial condition $\xi^{(n)}$. Let $\alpha_i^{(n)}(t) := \alpha(t, \frac{i}{n}, X_i^{(n)}(t))$ and $\widetilde{\alpha}_i^{(n)}(t) := \alpha(t, u, \widetilde{X}_u^{(n)}(t))$. It is obvious that for each realization $\boldsymbol{u}^{(n)}$ of $\boldsymbol{u}^{(n)}$ $\zeta_{ij}^{(n)}(\boldsymbol{u}^{(n)})$ satisfies regularity Assumption 6.5 with $G_n(\boldsymbol{u}^{(n)})$. Then, by an intermediate product in the proof of Theorem 6.10, we have

$$\frac{1}{n}\sum_{i=1}^n \mathbb{E}\|X_i^{(n)}(\boldsymbol{u}^{(n)}) - \widetilde{X}^{(n)}_{\frac{i}{n}}(\boldsymbol{u}^{(n)})\|^2_{\mathbb{S}^2_T} \to 0.$$

By Theorem 6.8, we have for each $\boldsymbol{u}^{(n)}$,

$$\int_I \mathbb{E}\|\widetilde{X}_v^{(n)}(\boldsymbol{u}^{(n)}) - X_v\|^2_{\mathbb{S}^2_T} dv \leqslant C\Big[\int_I \mathbb{E}|\xi_v^{(n)} - \xi_v|^2 dv + \|G - G_n(\boldsymbol{u}^{(n)})\|_\Box\Big].$$

Thus, we can recover the results in Theorem 6.10 with a stochastic version of $G_n(\boldsymbol{u}^{(n)})$.

We condition on the sampled sequence of step graphon $G_n(\boldsymbol{U}^{(n)})(u,v)$. Let $\Delta_i^{(n),1}(\alpha, \boldsymbol{u}^{(n)})$ be defined as following

$$\begin{aligned}\Delta_i^{(n),1}(\alpha, \boldsymbol{u}^{(n)}) := \mathbb{E}\Big[&\int_0^T f(t, X_i^{(n),\alpha,-i}(t), M_i^{(n),-i}(t), \alpha(t))dt + g(X_i^{(n),\alpha,-i}(T), M_i^{(n),-i}(T))\Big|\boldsymbol{u}^{(n)}\Big] \\ &-\mathbb{E}\Big[\int_0^T f(t, X^{\star,\alpha,-\frac{i}{n}}_{u_i^{(n)}}(t), \Lambda\mu_t^{\alpha,-\frac{i}{n}}(u_i^{(n)}), \alpha(t))dt + g(X^{\star,\alpha,-\frac{i}{n}}_{u_i^{(n)}}(T), \Lambda\mu_T^{\alpha,-\frac{i}{n}}(u_i^{(n)}))\Big],\end{aligned}$$

and similarly define $\Delta_i^{(n),3}(\alpha, \boldsymbol{u}^{(n)})$. We have similarly as in the continuous graphon case,

$$\frac{1}{n}\sum_{i=1}^n \Delta_i^{(n),1}(\alpha, \boldsymbol{u}^{(n)}) \leqslant C\Big[\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n(\boldsymbol{U}^{(n)})\|_\Box\Big] + o(1),$$

and

$$\frac{1}{n}\sum_{i=1}^n \Delta_i^{(n),1}(\alpha, \boldsymbol{u}^{(n)}) \leqslant C\Big[\frac{1}{n}\sum_{i=1}^n \mathbb{E}|\xi_i^{(n)} - \xi_{\frac{i}{n}}|^2 + \|G - G_n(\boldsymbol{U}^{(n)})\|_\Box\Big] + o(1).$$

Further, by [66, Theorem 2.14], $G_n(\boldsymbol{u}^{(n)})$ converges to G in probability in cut norm. Notice that $G_n(\boldsymbol{u}^{(n)})$ is bounded, so we have

$$\mathbb{E}\|G - G_n(\boldsymbol{U}^{(n)})\|_\Box \to 0.$$

Combining the above two results, we can conclude the proof of Theorem 6.20. □

6.7 Concluding remarks

We study stochastic graphon mean field games with jumps in the Markovian framework and approximate the Nash equilibria of finite games with heterogeneous interactions using the graphon equilibria as benchmarks. Notice that we are not able to solve the graphon equilibrium yet in the general case, although in some very specific cases we can. For classical mean field games, the solution can be characterized from either the PDE viewpoint, leading to a coupled Hamilton-Jacobi-Bellman equation and Kolmogorov-Fokker-Plank equation, or from a probabilistic point of view. The extensions to the graphon mean field framework will be very interesting and challenging.

The study of games based on more complex systems could be a direction for future work. As mentioned in the conclusion of Chapter 5, incorporating heterogeneous interactions that depend on the underlying network structure would be very interesting. The generalization to the non-Markovian framework is also valuable. In this work, we restrict the study to Markovian feedback controls. Usually, the type of control varies from model to model. An important extension can be related to controls with more general forms, particularly, the optimal stopping problems for graphon mean field games, which would absolutely attract a lot of attention. Another interesting direction would be generalise the strength of interactions ζ^n, and consider it evolves over time. We could associate it with a (controlled) process, similar as that in [61] or include it in the dynamics of the controlled state process as in [52]. The study of concentration results regarding the equilibrium approximation would also have significant application interest.

Bibliography

[1] Daron Acemoglu, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608, 2015.

[2] Viral V Acharya, Lasse H Pedersen, Thomas Philippon, and Matthew Richardson. Measuring systemic risk. *The Review of Financial Studies*, 30(1):2–47, 2017.

[3] Zoé Agathe-Nerine. Multivariate hawkes processes on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 152:86–148, 2022.

[4] Hansjörg Albrecher and Jevgenijs Ivanovs. Strikingly simple identities relating exit problems for lévy processes under continuous and poisson observations. *Stochastic Processes and their Applications*, 127(2):643–656, 2017.

[5] Hansjörg Albrecher, Jevgenijs Ivanovs, and Xiaowen Zhou. Exit identities for lévy processes observed at poisson arrival times. 2016.

[6] Charalambos D Aliprantis and Kim C Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer Science & Business Media, 2007.

[7] Franklin Allen and Douglas Gale. Financial contagion. *Journal of political economy*, 108(1):1–33, 2000.

[8] H. Amini. Bootstrap percolation in living neural networks. *Journal of Statistcal Physics*, 141:459–475, 2010.

[9] H. Amini and N. Fountoulakis. Bootstrap percolation in power-law random graphs. *Journal of Statistical Physics*, 155:72–92, 2014.

[10] Hamed Amini. Bootstrap percolation and diffusion in random graphs with given vertex degrees. *Electronic Journal of Combinatorics, 17: R25*, 2010.

[11] Hamed Amini. Bootstrap percolation in living neural networks. *Journal of Statistical Physics*, 141:459–475, 2010.

[12] Hamed Amini. Contagion risks and security investment in directed networks. *Mathematics and Financial Economics*, 17(2):247–283, 2023.

[13] Hamed Amini, Erhan Bayraktar, and Suman Chakraborty. A central limit theorem for diffusion in sparse random graphs. *Journal of Statistical Physics*, 190(3):57, 2023.

[14] Hamed Amini, Zhongyuan Cao, Andreea Minca, and Agnès Sulem. Ruin probabilities for risk processes in stochastic networks. *Available at SSRN 4355988*, 2022.

[15] Hamed Amini, Zhongyuan Cao, and Agnès Sulem. The default cascade process in stochastic financial networks. *Available at SSRN 4020598*, 2021.

[16] Hamed Amini, Zhongyuan Cao, and Agnès Sulem. Fire sales, default cascades and complex financial networks. *Available at SSRN 3935450 and in revision for Mathematical Finance*, 2021.

[17] Hamed Amini, Zhongyuan Cao, and Agnès Sulem. Limit theorems for default contagion and systemic risk. *Available at SSRN 3811107 and in revision for Mathematics of Operations Research*, 2021.

[18] Hamed Amini, Zhongyuan Cao, and Agnès Sulem. Graphon mean-field backward stochastic differential equations with jumps and associated dynamic risk measures. *Available at SSRN 4162616*, 2022.

[19] Hamed Amini, Zhongyuan Cao, and Agnès Sulem. Stochastic graphon mean field games with jumps and approximate nash equilibria. *Available at SSRN 4412999*, 2023.

[20] Hamed Amini, Rama Cont, and Andreea Minca. Resilience to contagion in financial networks. *Mathematical Finance*, 26(2):329–365, 2016.

[21] Hamed Amini and Zachary Feinstein. Optimal network compression. *arXiv preprint arXiv:2008.08733*, 2020.

[22] Hamed Amini, Damir Filipović, and Andreea Minca. Uniqueness of equilibrium in a payment system with liquidation costs. *Operations Research Letters*, 44(1):1–5, 2016.

[23] Hamed Amini, Nikolaos Fountoulakis, and Konstantinos Panagiotou. Bootstrap percolation in inhomogeneous random graphs. *arXiv preprint arXiv:1402.2815*, 2014.

[24] Hamed Amini and Andreea Minca. Inhomogeneous financial networks and contagious links. *Operations Research*, 64(5):1109–1120, 2016.

[25] Hamed Amini and Andreea Minca. Clearing financial networks: Impact on equilibrium asset prices and seniority of claims. In *Pushing the Boundaries: Frontiers in Impactful OR/OM Research*, pages 154–175. INFORMS, 2020.

[26] Hamed Amini, Andreea Minca, and Agnès Sulem. Control of interbank contagion under partial information. *SIAM Journal on Financial Mathematics*, 6(1):1195–1219, 2015.

[27] Hamed Amini, Andreea Minca, and Agnès Sulem. Optimal equity infusions in interbank networks. *Journal of Financial stability*, 31:1–17, 2017.

[28] Hamed Amini, Andreea Minca, and Agnès Sulem. A dynamic contagion risk model with recovery features. *Mathematics of Operations Research*, 47(2):1412–1442, 2021.

[29] Hamed Amini, Andreea Minca, and Agnès Sulem. A dynamic contagion risk model with recovery features. *Mathematics of Operations Research*, 47(2):1412–1442, 2022.

[30] Kartik Anand, Ben Craig, and Goetz Von Peter. Filling in the blanks: Network structure and interbank contagion. *Quantitative Finance*, 15(4):625–636, 2015.

[31] E Sparre Andersen. On the collective theory of risk in case of contagion between claims. *Bulletin of the Institute of Mathematics and its Applications*, 12(2):275–279, 1957.

[32] Andrea Angiuli, Nils Detering, Jean-Pierre Fouque, Mathieu Laurière, and Jimin Lin. Reinforcement learning for intra-and-inter-bank borrowing and lending mean field control game. In *Proceedings of the Third ACM International Conference on AI in Finance*, pages 369–376, 2022.

[33] David Applebaum. Nonlinear Markov processes and kinetic equations (cambridge tracts in mathematics 182) by Vassili N. Kolokoltsov. *Bulletin of the London Mathematical Society*, 43(6):1245–1247, 2011.

[34] Roc Armenter and Miklós Koren. A balls-and-bins model of trade. *American Economic Review*, 104(7):2127–51, 2014.

[35] Søren Asmussen and Hansjörg Albrecher. *Ruin Probabilities*. WORLD SCIENTIFIC, 2nd edition, 2010.

[36] Alexander Aurell, René Carmona, and Mathieu Laurière. Stochastic graphon games: Ii. the linear-quadratic case. *Applied Mathematics & Optimization*, 85(3):39, 2022.

[37] Florin Avram and Andreea Minca. On the management of central branch risk networks. *Advances in Applied probability, in print*, 2016.

[38] Florin Avram and Andreea Minca. On the central management of risk networks. *Advances in Applied Probability*, 49(1):221–237, 2017.

[39] Florin Avram, Zbigniew Palmowski, and Martijn Pistorius. A two-dimensional ruin problem on the positive quadrant. *Insurance: Mathematics and Economics*, 42(1):227–234, 2008.

[40] Florin Avram, Zbigniew Palmowski, and Martijn R. Pistorius. On the optimal dividend problem for a spectrally negative lévy process. *The Annals of Applied Probability*, 17(1):156–180, 2007.

[41] Florin Avram, Zbigniew Palmowski, and Martijn R Pistorius. Exit problem of a two-dimensional risk process from the quadrant: exact and asymptotic results. *The Annals of Applied Probability*, 18(6):2421–2449, 2008.

[42] Pablo Azcue and Nora Muler. Optimal reinsurance and dividend distribution policies in the cramér-lundberg model. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 15(2):261–308, 2005.

[43] Ana Babus. The formation of financial networks. *The RAND Journal of Economics*, 47(2):239–272, 2016.

[44] J. Balogh and B. G. Pittel. Bootstrap percolation on the random regular graph. *Random Structures & Algorithms*, 30(1-2):257–286, 2007.

[45] Pauline Barrieu and Nicole El Karoui. Optimal derivatives design under dynamic risk measures. *Contemporary Mathematics*, 351:13–26, 2004.

[46] Stefano Battiston, Guido Caldarelli, Robert M May, Tarik Roukny, and Joseph E Stiglitz. The price of complexity in financial networks. *Proceedings of the National Academy of Sciences*, 113(36):10031–10036, 2016.

[47] Erhan Bayraktar, Suman Chakraborty, and Ruoyu Wu. Graphon mean field systems. *to appear in The Annals of Applied Probability*, 2023.

[48] Erhan Bayraktar, Suman Chakraborty, and Xin Zhang. K-core in percolated dense graph sequences. *arXiv preprint arXiv:2012.09730*, 2020.

[49] Erhan Bayraktar, Gaoyue Guo, Wenpin Tang, and Yuming Zhang. Mckean-vlasov equations involving hitting times: blow-ups and global solvability. *arXiv preprint arXiv:2010.14646, to appear in The Annals of Applied Probability*, 2020.

[50] Erhan Bayraktar, Gaoyue Guo, Wenpin Tang, and Yuming Paul Zhang. Systemic robustness: a mean-field particle system approach. *arXiv preprint arXiv:2212.08518*, 2022.

[51] Erhan Bayraktar and Donghan Kim. Concentration of measure for graphon particle system. *arXiv preprint arXiv:2205.04951*, 2022.

[52] Erhan Bayraktar and Ruoyu Wu. Mean field interaction on random graphs with dynamically changing multi-color edges. *Stochastic Processes and their Applications*, 141:197–244, 2021.

[53] Erhan Bayraktar and Ruoyu Wu. Stationarity and uniform in time convergence for the graphon particle system. *Stochastic Processes and their Applications*, 150:532–568, 2022.

[54] Erhan Bayraktar and Ruoyu Wu. Graphon particle system: Uniform-in-time concentration bounds. *Stochastic Processes and their Applications*, 156:196–225, 2023.

[55] Erhan Bayraktar, Ruoyu Wu, and Xin Zhang. Propagation of chaos of forward-backward stochastic differential equations with graphon interactions. *Applied Mathematics and Optimization*, 88(1), August 2023.

[56] Anita Behme, Claudia Klüppelberg, and Gesine Reinert. Ruin probabilities for risk processes in a bipartite network. *Stochastic Models*, 36(4):548–573, 2020.

[57] Edward A. Bender and E. Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Comb. Theory, Ser. A*, 24:296–307, 1978.

[58] Sylvain Benoit, Jean-Edouard Colliard, Christophe Hurlin, and Christophe Pérignon. Where the risks lie: A survey on systemic risk. *Review of Finance*, 21(1):109–152, 2017.

[59] Amel Bentata and Rama Cont. Mimicking the marginal distributions of a semimartingale. Revision: 2011, 2009.

[60] Gianmarco Bet, Fabio Coppini, and Francesca R Nardi. Weakly interacting oscillators on dense random graphs. *arXiv preprint arXiv:2006.07670*, 2020.

[61] Shankar Bhamidi, Amarjit Budhiraja, and Ruoyu Wu. Weakly interacting particle systems on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 129(6):2174–2206, 2019.

[62] Maxim Bichuch and Zachary Feinstein. Optimization of fire sales and borrowing in systemic risk. *SIAM Journal on Financial Mathematics*, 10(1):68–88, 2019.

[63] P. Billingsley, Karreman Mathematics Research Collection, W.A. Shewhart, Wiley series in probability, mathematical statistics, and S.S. Wilks. *Convergence of Probability Measures*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1968.

[64] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

[65] B. Bollobas. *Random Graphs*. Cambridge University Press, 2001.

[66] Christian Borgs, Jennifer Chayes, Henry Cohn, and Yufei Zhao. An l^p theory of sparse graph convergence i: Limits, sparse random graph models, and power law distributions. *Transactions of the American Mathematical Society*, 372(5):3019–3062, 2019.

[67] Christian Borgs, Jennifer T Chayes, László Lovász, Vera T Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs i: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801–1851, 2008.

[68] Christian Borgs, Jennifer T Chayes, László Lovász, Vera T Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs ii. multiway cuts and statistical physics. *Annals of Mathematics*, pages 151–219, 2012.

[69] Michael Boss, Helmut Elsinger, Martin Summer, and Stefan Thurner. Network topology of the interbank market. *Quantitative Finance*, 4(6):677–684, 2004.

[70] Yann Braouezec and Lakshithe Wagalath. Strategic fire-sales and price-mediated contagion in the banking system. *European Journal of Operational Research*, 274(3):1180–1197, 2019.

[71] Gerard Brunick and Steven Shreve. Mimicking an itô process by a solution of a stochastic differential equation. *The Annals of Applied Probability*, 23(4):1584–1628, 2013.

[72] Rainer Buckdahn, Boualem Djehiche, Juan Li, and Shige Peng. Mean-field backward stochastic differential equations: a limit approach. *The Annals of Probability*, 37(4):1524–1565, 2009.

[73] Rainer Buckdahn, Juan Li, and Shige Peng. Mean-field backward stochastic differential equations and related partial differential equations. *Stochastic Processes and Their Applications*, 119(10):3133–3154, 2009.

[74] Amarjit Budhiraja, Paul Dupuis, Markus Fischer, and Kavita Ramanan. Limits of relative entropies associated with weakly interacting particle systems. *Electron. J. Probab*, 20(80):1–22, 2015.

[75] Amarjit Budhiraja and Ruoyu Wu. Some fluctuation results for weakly interacting multi-type particle systems. *Stochastic Processes and their Applications*, 126(8):2253–2296, 2016.

[76] Ricardo J Caballero and Alp Simsek. Fire sales in a model of complexity. *The Journal of Finance*, 68(6):2549–2587, 2013.

[77] Antonio Cabrales, Piero Gottardi, and Fernando Vega-Redondo. Risk sharing and contagion in networks. *The Review of Financial Studies*, 30(9):3086–3127, 2017.

[78] Peter E Caines, Daniel WC HO, Minyi Huang, Jiamin Jian, and Qingshuo Song. On the graphon mean field game equations: Individual agent affine dynamics and mean field dependent performance functions. *arXiv preprint arXiv:2009.12144*, 2020.

[79] Peter E Caines and Minyi Huang. Graphon mean field games and their equations. *SIAM Journal on Control and Optimization*, 59(6):4373–4399, 2021.

[80] Agostino Capponi and Martin Larsson. Price contagion through balance sheet linkages. *The Review of Asset Pricing Studies*, 5(2):227–253, 2015.

[81] Agostino Capponi, Xu Sun, and David D Yao. A dynamic network model of interbank lending—systemic risk and liquidity provisioning. *Mathematics of Operations Research*, 45(3):1127–1152, 2020.

[82] René Carmona, Daniel B Cooney, Christy V Graves, and Mathieu Lauriere. Stochastic graphon games: I. the static case. *Mathematics of Operations Research*, 47(1):750–778, 2022.

[83] René Carmona, François Delarue, et al. *Probabilistic theory of mean field games with applications I-II*. Springer, 2018.

[84] René Carmona, François Delarue, and Daniel Lacker. Mean field games with common noise. *Annals of Probability*, 44(6):3740–3803, 2016.

[85] Rene A Carmona, Jean Pierre Fouque, and Li Hsien Sun. Mean field games and systemic risk. *Communications in Mathematical Sciences*, 13(4):911–933, 2015.

[86] J Chalupa, P L Leath, and G R Reich. Bootstrap percolation on a bethe lattice. *Journal of Physics C: Solid State Physics*, 12(1):L31, jan 1979.

[87] Chen Chen, Garud Iyengar, and Ciamac C. Moallemi. An axiomatic approach to systemic risk. *Management Science*, 59(6):1373–1388, 2013.

[88] Nan Chen, Xin Liu, and David D Yao. An optimization view of financial systemic risk modeling: Network effect and market liquidity effect. *Operations research*, 64(5):1089–1108, 2016.

[89] Rui Chen, Roxana Dumitrescu, Andreea Minca, and Agnès Sulem. Mean-field BSDEs with jumps and dual representation for global risk measures. *Probability, Uncertainty and Quantitative Risk*, 8(1):33-52,2023.

[90] Sergey Chernenko and Adi Sunderam. Do fire sales create externalities? *Journal of Financial Economics*, 135(3):602–628, 2020.

[91] Matteo Chinazzi and Giorgio Fagiolo. Systemic risk, contagion, and financial networks: A survey. *Institute of Economics, Scuola Superiore Sant'Anna, Laboratory of Economics and Management (LEM) Working Paper Series*, (2013/08), 2015.

[92] Rodrigo Cifuentes, Gianluigi Ferrucci, and Hyun Song Shin. Liquidity risk and contagion. *Journal of the European Economic Association*, 3(2-3):556–566, 2005.

[93] Rama Cont, Amal Moussa, et al. Network structure and systemic risk in banking systems. *Edson Bastos e, Network Structure and Systemic Risk in Banking Systems (December 1, 2010)*, 2010.

[94] Rama Cont and Eric Schaanning. Fire sales, indirect contagion and systemic stress testing. *Indirect Contagion and Systemic Stress Testing (June 13, 2017)*, 2017.

[95] Rama Cont and Eric Schaanning. Monitoring indirect contagion. *Journal of Banking & Finance*, 104:85–102, 2019.

[96] Rama Cont and Lakshithe Wagalath. Fire sales forensics: measuring endogenous risk. *Mathematical finance*, 26(4):835–866, 2016.

[97] Fabio Coppini. Long time dynamics for interacting oscillators on graphs. *The Annals of Applied Probability*, 32(1):360–391, 2022.

[98] Emilie Coupechoux and Marc Lelarge. How clustering affects epidemics in random networks. *Advances in Applied Probability*, 46(4):985–1008, 2014.

[99] Ben Craig and Goetz Von Peter. Interbank tiering and money center banks. *Journal of Financial Intermediation*, 23(3):322–347, 2014.

[100] Harald Cramer. *Collective risk theory.* Skandia ins. Company, 1955.

[101] Kai Cui and Heinz Koeppl. Learning graphon mean field games and approximate nash equilibria. *arXiv preprint arXiv:2112.01280*, 2021.

[102] François Delarue. Mean field games: A toy model on an erdös-renyi graph. *ESAIM: Proceedings and Surveys*, 60:1–26, 2017.

[103] Freddy Delbaen and Jean Haezendonck. Classical risk theory in an economic environment. *Insurance: Mathematics and Economics*, 6(2):85–116, 1987.

[104] Marco D'Errico and Tarik Roukny. Compressing over-the-counter markets. *Available at SSRN 2962575*, 2019.

[105] Nils Detering, Thilo Meyer-Brandis, Konstantinos Panagiotou, and Daniel Ritter. Managing default contagion in inhomogeneous financial networks. *SIAM Journal on Financial Mathematics*, 10(2):578–614, 2019.

[106] Nils Detering, Thilo Meyer-Brandis, Konstantinos Panagiotou, and Daniel Ritter. An integrated model for fire sales and default contagion. *Mathematics and Financial Economics*, 15(1):59–101, 2021.

[107] Nils Detering, Thilo Meyer-Brandis, Konstantinos Panagiotou, and Daniel Ritter. Suffocating fire sales. *SIAM Journal on Financial Mathematics*, 13(1):70–108, 2022.

[108] Fernando Duarte and Thomas M Eisenbach. Fire-sale spillovers and systemic risk. *The Journal of Finance*, 76(3):1251–1294, 2021.

[109] Richard M. Dudley, Eva Giné, and Joel C. Zinn. Uniform and universal glivenko-cantelli classes. *Journal of Theoretical Probability*, 4:485–510, 1991.

[110] Paul Dupuis and Georgi S Medvedev. The large deviation principle for interacting dynamical systems on random graphs. *Communications in Mathematical Physics*, 390(2):545–575, 2022.

[111] Larry Eisenberg and Thomas H Noe. Systemic risk in financial systems. *Management Science*, 47(2):236–249, 2001.

[112] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems.* SIAM, 1999.

[113] Nicole El Karoui, Du Huu Nguyen, and Jeanblanc-Picqué Monique. Compactification methods in the control of degenerate diffusions: existence of an optimal control. *Stochastics: an international journal of probability and stochastic processes*, 20(3):169–219, 1987.

[114] Matthew Elliott, Benjamin Golub, and Matthew O Jackson. Financial networks and contagion. *American Economic Review*, 104(10):3115–53, 2014.

[115] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media, 2013.

[116] Selman Erol and Rakesh Vohra. Network formation and systemic risk. *European Economic Review*, 148:104213, 2022.

[117] Zachary Feinstein. Financial contagion and asset liquidation strategies. *Operations Research Letters*, 45(2):109–114, 2017.

[118] Zachary Feinstein, Birgit Rudloff, and Stefan Weber. Measures of systemic risk. *SIAM Journal on Financial Mathematics*, 8(1):672–708, 2017.

[119] Zachary Feinstein and Andreas Søjmark. Dynamic default contagion in heterogeneous interbank systems. *SIAM Journal on Financial Mathematics*, 12(4):SC83–SC97, 2021.

[120] Yichen Feng, Jean-Pierre Fouque, and Tomoyuki Ichiba. Linear-quadratic stochastic differential games on directed chain networks. *arXiv preprint arXiv:2003.08840*, 2020.

[121] Hans Föllmer. Random economies with many interacting agents. *Journal of mathematical economics*, 1(1):51–62, 1974.

[122] Daniel Fricke and Thomas Lux. Core–periphery structure in the overnight money market: evidence from the e-mid trading platform. *Computational Economics*, 45(3):359–395, 2015.

[123] Marco Frittelli and E Rosazza Gianin. Dynamic convex risk measures. *Risk measures for the 21st century*, pages 227–248, 2004.

[124] Prasanna Gai and Sujit Kapadia. Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466(2120):2401–2423, 2010.

[125] Andrea Galeotti, Benjamin Golub, and Sanjeev Goyal. Targeting interventions in networks. *Econometrica*, 88(6):2445–2471, 2020.

[126] Axel Gandy and Luitgard AM Veraart. A bayesian methodology for systemic risk assessment in financial networks. *Management Science*, 63(12):4428–4446, 2017.

[127] Shuang Gao, Rinel Foguen Tchuendom, and Peter E Caines. Linear quadratic graphon field games. *arXiv preprint arXiv:2006.03964*, 2020.

[128] Hans U Gerber. *Entscheidungskriterien für den zusammengesetzten Poisson-Prozess*. PhD thesis, ETH Zurich, 1969.

[129] Kay Giesecke and Stefan Weber. Cyclical correlations, credit contagion, and portfolio losses. *Journal of Banking & Finance*, 28(12):3009–3036, 2004.

[130] Kay Giesecke and Stefan Weber. Credit contagion and aggregate losses. *Journal of Economic Dynamics and Control*, 30(5):741–767, 2006.

[131] Paul Glasserman and H Peyton Young. How likely is contagion in financial networks? *Journal of Banking & Finance*, 50:383–399, 2015.

[132] Paul Glasserman and H Peyton Young. Contagion in financial networks. *Journal of Economic Literature*, 54(3):779–831, 2016.

[133] Robin Greenwood, Augustin Landier, and David Thesmar. Vulnerable banks. *Journal of Financial Economics*, 115(3):471–485, 2015.

[134] Haotian Gu, Xin Guo, Xiaoli Wei, and Renyuan Xu. Mean-field multi-agent reinforcement learning: A decentralized network approach. *arXiv preprint arXiv:2108.02731*, 2021.

[135] Ben Hambly, Sean Ledger, and Andreas Søjmark. A mckean–vlasov equation with positive feedback and blow-ups. *The Annals of Applied Probability*, 29(4):2338–2373, 2019.

[136] Ben Hambly and Andreas Søjmark. An spde model for systemic risk with endogenous contagion. *Finance and Stochastics*, 23(3):535–594, 2019.

[137] Tao Hao and Juan Li. Mean-field SDEs with jumps and nonlocal integral-PDEs. *Nonlinear Differential Equations and Applications NoDEA*, 23(2):1–51, 2016.

[138] UG Haussmann and JP Lepeltier. On the existence of optimal controls. *SIAM Journal on Control and Optimization*, 28(4):851–902, 1990.

[139] Remco van der Hofstad. *Random Graphs and Complex Networks*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2016.

[140] Yuanquan Hu, Xiaoli Wei, Junji Yan, and Hengxi Zhang. Graphon mean-field control for cooperative multi-agent reinforcement learning. *arXiv preprint arXiv:2209.04808*, 2022.

[141] Minyi Huang, Peter E Caines, and Roland P Malhamé. Large-population cost-coupled lqg problems with nonuniform agents: individual-mass behavior and decentralized ε-nash equilibria. *IEEE transactions on automatic control*, 52(9):1560–1571, 2007.

[142] Minyi Huang, Roland P Malhamé, and Peter E Caines. 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.

[143] Thomas R Hurd. *Contagion! Systemic risk in financial networks*. Springer, 2016.

[144] Giulia Iori, Rosario N Mantegna, Luca Marotta, Salvatore Micciche, James Porter, and Michele Tumminello. Networked relationships in the e-mid interbank market: A trading model with memory. *Journal of Economic Dynamics and Control*, 50:98–116, 2015.

[145] Matthew O Jackson and Agathe Pernoud. Credit freezes, equilibrium multiplicity, and optimal bailouts in financial networks. *arXiv preprint arXiv:2012.12861*, 2020.

[146] Matthew O Jackson and Agathe Pernoud. Systemic risk in financial networks: A survey. *Available at SSRN*, 2020.

[147] Jean Jacod and Albert Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.

[148] S. Janson, T. Łuczak, T. Turova, and T. Vallier. Bootstrap percolation on the random graph $G_{n,p}$. *The Annals of Applied Probability*, 22(5):1989–2047, 2012.

[149] Svante Janson. The probability that a random multigraph is simple. *Combinatorics, Probability and Computing*, 18(1-2):205–225, 2009.

[150] Svante Janson et al. On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*, 14:86–118, 2009.

[151] Svante Janson and Malwina J. Luczak. A simple solution to the k-core problem. *Random Struct. Algorithms*, 30(1–2):50–62, jan 2007.

[152] Svante Janson, Malwina J Luczak, et al. Asymptotic normality of the k-core in random graphs. *The annals of applied probability*, 18(3):1085–1137, 2008.

[153] Svante Janson, Malwina J. Luczak, and Peter Windridge. Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures & Algorithms*, 45(4):726–763, 2014.

[154] Mark Kac. *Probability and related topics in physical sciences*, volume 1. American Mathematical Soc., 1959.

[155] Dmitry Kaliuzhnyi-Verbovetskyi and Georgi S Medvedev. The mean field equation for the kuramoto model on graph sequences with non-lipschitz limit. *SIAM Journal on Mathematical Analysis*, 50(3):2441–2465, 2018.

[156] Olav Kallenberg. *Foundations of modern probability*, volume 2. Springer, 1997.

[157] Jon Kleinberg. Cascading behavior in networks: Algorithmic and economic issues. *Algorithmic game theory*, 24:613–632, 2007.

[158] Oliver Kley, Claudia Klüppelberg, and Gesine Reinert. Risk in a large claims insurance market with bipartite graph structure. *Operations Research*, 64(5):1159–1176, 2016.

[159] Eduard Kromer, Ludger Overbeck, and Katrin Zilch. Systemic risk measures on general probability spaces. *Mathematical Methods of Operations Research*, 84(2):323–357, 2016.

[160] Daniel Lacker. Mean field games via controlled martingale problems: Existence of markovian equilibria. *Stochastic Processes and their Applications*, 125(7):2856–2894, 2015.

[161] Daniel Lacker and Agathe Soret. A case study on stochastic games on large graphs in mean field and sparse regimes. *Mathematics of Operations Research*, 47(2):1530–1565, 2022.

[162] Daniel Lacker and Agathe Soret. A label-state formulation of stochastic graphon games and approximate equilibria on large networks. *Mathematics of Operations Research*, 2022.

[163] Jean-Michel Lasry and Pierre-Louis Lions. Mean Field Games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[164] Claude Lefèvre and Stéphane Loisel. On Finite-Time Ruin Probabilities for Classical Risk Models. *Scandinavian Actuarial Journal*, 2008(1):41–60, January 2008.

[165] Marc Lelarge. Diffusion and cascading behavior in random networks. *Games and Economic Behavior*, 75(2):752–775, 2012.

[166] Dan Li and Norman Schürhoff. Dealer networks. *The Journal of Finance*, 74(1):91–144, 2019.

[167] Juan Li. Mean-field forward and backward SDEs with jumps and associated nonlocal quasi-linear integral-PDEs. *Stochastic Processes and their Applications*, 128(9):3118–3180, 2018.

[168] Juan Li and Hui Min. Controlled mean-field backward stochastic differential equations with jumps involving the value function. *Journal of Systems Science and Complexity*, 29(5):1238–1268, 2016.

[169] Ronnie L Loeffen. On optimality of the barrier strategy in de finetti's dividend problem for spectrally negative lévy processes. *The Annals of Applied Probability*, pages 1669–1680, 2008.

[170] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[171] Fillip Lundberg. *Approximerad framställning av sannolikhetsfunktionen. aterförsäkring av kollektivrisker. Akad.* PhD thesis, Almqvist och Wiksell Uppsala, Sweden, 1903.

[172] Henry P McKean. Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57, 1967.

[173] Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press, 2015.

[174] Stephen Morris. Contagion. *The Review of Economic Studies*, 67(1):57–78, 2000.

[175] Stephen Morris and Hyun Song Shin. Illiquidity component of credit risk. Technical report, working paper, Princeton University, 2009.

[176] Jeannette Müller. Interbank credit lines as a channel of contagion. *Journal of Financial Services Research*, 29(1):37–60, 2006.

[177] Sergey Nadtochiy and Mykhaylo Shkolnikov. Particle systems with singular interaction through hitting times: Application in systemic risk modeling. *Annals of Applied Probability*, 29(1):89–129, 2019.

[178] Sergey Nadtochiy and Mykhaylo Shkolnikov. Mean field systems on networks, with singular interaction through hitting times. *Annals of Probability*, 48(3), 2020.

[179] Erlend Nier, Jing Yang, Tanju Yorulmazer, and Amadeo Alentorn. Network models and financial stability. *Journal of Economic Dynamics and Control*, 31(6):2033–2060, 2007.

[180] Bernt Øksendal and Agnes Sulem. *Stochastic Control of jump diffusions*. Springer, 2005.

[181] Roberto I Oliveira and Guilherme H Reis. Interacting diffusions on random graphs with diverging average degrees: Hydrodynamics and large deviations. *Journal of Statistical Physics*, 176(5):1057–1087, 2019.

[182] Francesca Parise and Asuman Ozdaglar. Graphon games: a statistical framework for network games and interventions. *Econometrica*, 91(1):191–225, 2023.

[183] Marie-Claire Quenez and Agnès Sulem. BSDEs with jumps, optimization and applications to dynamic risk measures. *Stochastic Processes and their Applications*, 123(8):3328–3357, 2013.

[184] L. C. G. Rogers and David Williams. *Diffusions, Markov Processes, and Martingales*, volume 1 of *Cambridge Mathematical Library*. Cambridge University Press, 2 edition, 2000.

[185] Leonard CG Rogers and Luitgard AM Veraart. Failure and rescue in an interbank network. *Management Science*, 59(4):882–898, 2013.

[186] Sheldon M Ross, John J Kelly, Roger J Sullivan, William James Perry, Donald Mercer, Ruth M Davis, Thomas Dell Washburn, Earl V Sager, Joseph B Boyce, and Vincent L Bristow. *Stochastic processes*, volume 2. Wiley New York, 1996.

[187] Tarik Roukny, Stefano Battiston, and Joseph E Stiglitz. Interconnectedness as a source of uncertainty in systemic risk. *Journal of Financial Stability*, 35:93–106, 2018.

[188] Hanspeter Schmidli. Optimisation in non-life insurance. *Stochastic models*, 22(4):689–722, 2006.

[189] Andrei Shleifer and Robert Vishny. Fire sales in finance and macroeconomics. *Journal of economic perspectives*, 25(1):29–48, 2011.

[190] Daniel W Stroock and SR Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233. Springer Science & Business Media, 1997.

[191] Shanjian Tang and Xunjing Li. Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM Journal on Control and Optimization*, 32(5):1447–1475, 1994.

[192] Adrian Tobias and Markus K Brunnermeier. Covar. *The American Economic Review*, 106(7):1705, 2016.

[193] Christian Upper and Andreas Worms. Estimating bilateral exposures in the german interbank market: Is there a danger of contagion? *European economic review*, 48(4):827–849, 2004.

[194] Remco van der Hofstad. *Random Graphs and Complex Networks*. Cambridge University Press, 2016.

[195] Remco van der Hofstad, Johan SH van Leeuwaarden, and Clara Stegehuis. Hierarchical configuration model. *arXiv preprint arXiv:1512.08397*, 2015.

[196] Luitgard A.M. Veraart. When does portfolio compression reduce systemic risk? *Mathematical Finance*, 32(3):727–778, 2022.

[197] Luitgard Anna Maria Veraart. When does portfolio compression reduce systemic risk? *Available at SSRN 3688495*, 2020.

[198] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[199] Daniel H Wagner. Survey of measurable selection theorems. *SIAM Journal on Control and Optimization*, 15(5):859–903, 1977.

[200] Duncan J Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771, 2002.

[201] Stefan Weber and Kerstin Weske. The joint impact of bankruptcy costs, fire sales and cross-holdings on systemic risk in financial networks. *Probability, Uncertainty and Quantitative Risk*, 2(1):1–38, 2017.

RÉSUMÉ

Cette thèse est divisée en deux parties. La première partie étudie la stabilité et le risque systémique de réseaux financiers complexes, soumis à des processus de contagion de défauts, et de ventes forcées. Nous prouvons des théorèmes limites de type loi des grands nombres et limite centrale sur la dynamique de contagion. Nous montrons comment quantifier le risque systémique d'un réseau financier en présence d'une perturbation externe et sous information partielle. Nous étudions ensuite les processus de risque multidimensionnels de Cramér-Lundberg où les agents, situés sur un grand réseau, subissent des pertes de la part de leurs voisins. Nous présentons enfin un cadre général abordable pour comprendre l'impact conjoint de liquidations et de cascades de défauts sur le risque systémique dans les réseaux financiers complexes. La deuxième partie de la thèse est consacrée à l'étude et le contrôle de systèmes interactifs de type graphon champ moyen. Le réseau financier est ici considéré comme un grand système interactif, ce qui établit un lien avec la théorie des jeux à champ moyen. La structure en champ moyen repose sur la structure de graphe sous-jacente du réseau, appelée champ moyen graphon. Nous commençons par une étude systématique des équations différentielles stochastiques rétrogrades (EDSR) avec sauts de type graphon champ moyen et ses mesures de risque dynamiques associées. Nous étudions ensuite des jeux stochastiques continus avec interactions non homogènes de type champ moyen sur de vastes réseaux et explorons leurs limites graphon champ moyen. Nous proposons des équilibres de Nash approximés pour les jeux finis sur les réseaux, utilisant les équilibres en champ moyen graphon associés comme référence.

MOTS CLÉS

Réseaux financiers, Contagion des défauts, Ventes forcées, Processus de risque, Théorèmes limites, Graphon champ moyen, Systèmes interactifs, Jeux stochastiques.

ABSTRACT

This thesis is divided in two parts. The first part considers the issues of stability and systemic risk in large complex financial networks, including the study of default contagion, fire sales and risk processes on networks. We first prove limit theorems (law of large numbers and central limit theorem types) for the contagion dynamics. We show how to quantify the systemic risk for a financial network under partial information facing an outside shock. Then we present a general tractable framework for understanding the joint impact of fire sales and default cascades on systemic risk in complex financial networks. We finally study risk processes on large financial systems, when agents, located on a large network, receive losses from their neighbors. The second part of the thesis focuses on graphon mean field interacting systems with jumps and graphon mean field games. Here, the financial network is seen as a large interacting system, with a graphon mean field structure depending on the underlying graph structure of the network. We first conduct a comprehensive study of graphon mean field backward stochastic differential equations (BSDEs) with jumps and associated global dynamic risk measures. We then study continuous stochastic games with heterogeneous mean field interactions on large networks and investigate their graphon limits. We provide approximate Nash equilibria for finite games with heterogeneous interactions, using their graphon equilibria as benchmarks.

KEYWORDS

Financial networks, Default contagion, Fire sales, Risk processes, Limit Theorems, Graphon mean field, Interacting systems, Stochastic games