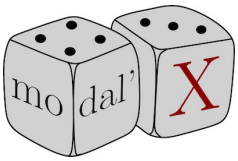

Université Paris Nanterre

École doctorale «Connaissance, langage modélisation» (ED 139)

Laboratoire Modal'X (EA 3454)

Mémoire présenté pour l'obtention du

Diplôme d'habilitation à diriger les recherches

Discipline : Mathématiques Appliquées

par

Hanene Mohamed

Modèles et techniques probabilistes autour de l'autopartage

Proposition de rapporteurs :

Carl Graham
Pascal Moyal
Marie Theret

Proposition de jury :

Laurent Decreusefond	(Examinateur)
Carl Graham	(Rapporteur)
Pascal Moyal	(Rapporteur)
Philippe Robert	(Examinateur)
Francesco Russo	(Examinateur)
Marie Théret	(Rapporteuse)
Viet Chi Tran	(Examinateur)
Amandine Veber	(Examinatrice)

Table des matières

Remerciements

Je tiens tout d'abord à remercier Philippe Soulier pour ses encouragements et sa bienveillance : il a su, sans me mettre la pression, me motiver afin que je me décide enfin à passer ce cap. Merci aussi à Marie Théret, ma référente, pour son intérêt, ses conseils (qui ont bien dépassé l'aspect scientifique) et le suivi impeccable de toute la procédure administrative. Je remercie aussi chaleureusement Carl Graham et Pascal Moyal d'avoir accepté la tâche ingrate de rapporter ce mémoire. Merci à tous les membres du jury d'avoir accepté de m'accompagner dans cette étape ultime de ce projet. C'est à la fois un honneur et un plaisir que vous me faites. Je remercie aussi mes collègues et anciens collègues de Modal'X, ce laboratoire de recherche où il fait bon travailler.

Dans ce mémoire, je présente des travaux que je n'ai pas réalisés seule. Un grand merci à mes co-auteurs : Alessia, Amaury, Bianca, Cédric, Christine, Hosam, Julien, Martin, Nicolas, Plinio, Rafik, Teodora, Toshio et Zhou. Merci à Christine Fricker de m'avoir fait découvrir le partage de véhicules comme une jolie application probabiliste qui a motivé nos nombreux travaux communs. Merci à Martin Trépanier qui nous a permis une collaboration précieuse avec l'opérateur canadien Communauto et donc l'opportunité d'encadrer des stages de recherche autour de cette thématique. J'ai aussi la chance de co-encadrer la thèse d'Alessia Rigonat dont l'énergie et l'enthousiasme font que travailler ensemble est toujours un réel plaisir.

Enfin, merci à mes proches pour leur soutien. Merci à ma sœur Jihène pour sa présence sans faille pendant les moments les plus durs et pour tout le reste. J'ai une pensée pour mes enfants Fehd, Yacine et Kenza, qui ont été témoins de l'intensité de l'exercice lors de la rédaction de ce manuscrit ou de la préparation de la soutenance. Ils ont été ma garde rapprochée essayant à leur échelle de me faire bénéficier de la meilleure ambiance afin que j'arrive à bout de cette épreuve. Je leur dis simplement que je les aime tant.

2025,
l'année de tous les défis :
Un passage en sixième,
un brevet,
un baccalauréat
et une HDR !

Avant-Propos

Ce document reprend les travaux de recherche que j'ai menés depuis ma thèse de doctorat. Les trois premiers chapitres en constituent une partie majeure, consacrée à un axe thématique dominant où je me suis interessée aux grands réseaux stochastiques d'un point de vue modélisation stochastique et analyse en temps long avec une approche champ-moyen. La motivation première de ces travaux de recherche est celle des systèmes de partage de véhicules, mon premier article sur le sujet date de 2012. Mes travaux sur le sujet forment un ensemble homogène dont une synthèse est présentée ici. J'ai tenu à valoriser la motivation de ces travaux car ils viennent répondre à des questions pratiques de dimensionnement de réseaux de partage de véhicules, d'analyse de performance ou de régulation du système. Un grand travail de modélisation résulte d'un échange fructueux avec l'opérateur Communauto à Montréal, mais aussi à Paris. Souvent des analyses de données ont guidé le choix des modèles proposés. Les techniques utilisées sont probabilistes relevant principalement de limite champ-moyen et de la théorie des files d'attente. D'autres outils tels que les techniques de renormalisation (limites fluides), couplage ou *averaging* stochastique ont servi à prouver des résultats de comportement en temps long de grands systèmes en interaction. Parmi les modèles présentés, je me suis attardée sur trois. D'abord le modèle homogène sans réservation (type Vélib'), qui me sert de modèle jouet pour illustrer l'approche champ-moyen ainsi que la mesure de performance du système. Le deuxième est le modèle de la double réservation (type Autolib') présentant des interactions fortes que l'approximation champ-moyen gomme, comme c'est le cas pour le célèbre modèle de Gibbens, Hunt et Kelly. Et enfin, le modèle de *free-floating*, les systèmes d'autopartage sans stations physiques, qui présente un principe d'*averaging* stochastique, une complication supplémentaire pour l'approche champ-moyen, avec une transition de phase et une analyse de performance toute à fait inattendue.

La deuxième partie, plus modeste et plus récente, comporte un chapitre, le dernier. C'est un retour vers des structures discrètes que j'ai eu l'occasion de découvrir lors de ma thèse ou encore mon séjour postdoctoral telles que les arbres aléatoires, les marches aléatoires ou encore les urnes. Il s'agit de l'arbre exponentiellement préférentiel, un arbre aléatoire où l'insertion d'un nouvel élément obéit à une règle de poids exponentiel. Malgré la définition simple de cette règle d'affinité, le comportement d'un tel arbre à grande taille s'avère riche. On obtient trois régimes, dont celui critique correspond à l'arbre récursif uniforme, bien connu. Pour chaque régime, trois phases apparaissent avec différents comportements : un Théorème Central Limit, des oscillations ou convergence. Les outils utilisés sont combinatoires ainsi que des techniques d'approximation de Poisson.

J'ai tenu à ce que ce mémoire soit aussi peu technique que possible, mettant l'accent sur les heuristiques plutôt que les preuves détaillées dans les articles listés (voir ma liste des publications à la fin de ce mémoire ou encore sur ma page web `https://mohamed.perso.math.cnrs.fr/`). À la fin de chaque chapitre, j'ai essayé d'aborder des questions ou pistes ouvertes, souvent des travaux en cours à différents niveaux d'avancement.

Chapitre 1

Modélisation probabiliste d'un système de vélopartage

Collaborations avec Julien Ancel, Plinio S. Dester, Christine Fricker et Nicolas Gast.

1.1 Motivation

Les systèmes de vélopartage (ou systèmes de vélos en libre-service) sont de plus en plus importants pour le transport urbain. Ils sont principalement utilisés pour les trajets courts. Historiquement, le premier système de vélopartage a été lancé à Copenhague en 1995. Mais le premier grand système à être déployé est celui de la ville de Paris, Vélib', lancé en juillet 2007 à grande échelle (environ 20 000 vélos et 1 500 stations). Depuis, des systèmes de vélopartage ont vu le jour dans les grandes villes du monde. Actuellement, plus de 400 villes sont équipées de tels systèmes. La popularité de ces systèmes de vélos en libre-service est à l'origine d'une activité de recherche récente qui vient répondre à des problématiques applicatives réelles.
Le concept d'un système de vélopartage est simple : un usager arrive à une station, prend un vélo, l'utilise pendant un certain temps et le rend ensuite à une autre station. Le manque de ressources est l'un des principaux problèmes. Il se produit lorsqu'un usager arrive à une station où il n'y a pas de vélo disponible, ou lorsqu'à la fin de son trajet, il arrive à une station où il n'y a pas de place libre et donc ne peut pas rendre son vélo. L'allocation des ressources, vélos et places libres, est un enjeu majeur pour l'opérateur afin d'offrir une alternative fiable aux autres modes de transport.
Un tel système doit répondre à la fois à la demande de vélos et de places libres. Cette demande est complexe, en temps (l'heure de la journée, le jour de la semaine, de la saison et de la météo) avec une pseudo-périodicité sur 24 h, mais aussi en espace (zones d'habitation ou de travail, gares, stations en altitude . . .). De plus, le système est stochastique en raison de l'aléa des arrivées aux stations, des couples origine-destination et de la longueur des trajets. Le manque de ressources génère également des choix aléatoires de la part des usagers qui doivent chercher une autre station.
L'approche classique pour étudier ces systèmes a été longtemps axée sur la recherche opérationnelle. Le rééquilibrage du système, c'est-à-dire redistribuer les vélos sur les stations, à travers la régulation par camions ou autres est l'une des questions importantes dressées. Cette approche ne tient pas compte de l'aléa dans le fonctionnement de tels systèmes. Quelques articles traitent de la stochasticité des systèmes de vélopartage, ou plus généralement des systèmes de partage de véhicules (voir [FL96], [GB02], [GX11], [FG16] et [FGM12] . . .) avec l'objectif d'obtenir un comportement asymptotique simplifié lorsque le système devient grand. Cette approximation est valable car ces systèmes sont de grande taille et peut établir des propriétés qualitatives et quantitatives du système. Elle permet aussi

d'aborder la problématique de la régulation du système en étudiant l'impact de différentes politiques incitatives et de proposer des scénarios à mettre en place par l'opérateur.
Mathématiquement, les systèmes de vélopartage ont d'abord été modélisés en tant que grands réseaux fermés, au sens des files d'attente, avec une capacité infinie, voir [GX10] et [FL96]. Il s'agissait de considérer la mesure invariante de l'état du système, explicite et de forme produit dans ce cas. Des asymptotiques peuvent être obtenues via l'analyse complexe (méthode du point de selle), voir [MV96], ou par des outils probabilistes, voir [FL96]. Ne tenant pas compte du fait que les stations ont des capacités finies, ces travaux négligent donc l'effet de saturation. Dans [FT17], une politique de blocage-reroutage, proposée par [EF98], permet de garder la forme produit de la mesure invariante. Le comportement asymptotique quand la taille du réseau devient grande est obtenu en utilisant le Théorème limite local. Néanmoins, dès que le modèle est modifié pour tenir compte d'une politique incitative ou d'une recherche locale en cas de blocage, la forme produit n'est plus vraie, ou à minima non connue. Ceci montre la fragilité de cette approche.
Une autre façon d'aborder le problème est de considérer la convergence de la mesure empirique, le système dynamique limite et son point d'équilibre. Cette approche champ moyen s'avère assez robuste, au fur et à mesure que des variantes du modèle de base sont étudiées pour tenir compte de vrais aspects applicatifs.

1.2 Limite champ-moyen

Le but de cette section est d'exhiber, sur un cas plus simple celui d'un système de vélopartage homogène, l'approche champ-moyen qui sera ensuite appliquée aux systèmes d'autopartage et faciliter ainsi la compréhension des modèles et techniques qui seront présentés dans les chapitres suivants.

Un modèle homogène

Considérons M vélos qui se déplacent entre N stations. M est du même ordre que N de telle façon que M/N tend vers une constate s quand N tend vers l'infini. Le ratio s est important pour l'opérateur car il représente le nombre moyen de vélos par station. C'est un des leviers d'ajustement permettant à l'opérateur d'agir sur les performances du système. Un grand ratio risque de saturer les stations et donc l'usager ne trouve pas de place libre à destination, un ratio faible pose le problème de l'absence de vélos disponibles dans les stations. Trouver le bon ratio s est ce qu'on appellera résoudre *le problème de dimensionnement*. Les paramètres du modèle ne dépendent pas des stations. Ainsi, toutes les stations ont une capacité finie c. Notons par λ le taux d'arrivée des usagers à une station donnée (λ est aussi le taux de départ des vélos d'une station donnée) et par $1/\mu$ la durée moyenne de trajet entre deux stations données. Toutes les durées inter-arrivées et de trajet ont des lois exponentielles et sont indépendantes entre elles.
On s'intéresse au processus $\left(X^N(t)\right) := \left(X_i^N(t), 1 \leqslant i \leqslant N\right)$ où $X_i^N(t)$ est le nombre de vélos dans la station i à l'instant t. Le processus de Markov $\left(X^N(t)\right)$ est irréductible sur l'espace d'états fini suivant :

$$\chi = \left\{ x = (x_i) \in \mathbb{N}^N, \forall\, 1 \leq i \leq N :\ x_i \leqslant c\,,\ \sum_{i=1}^{N} x_i \leqslant M \right\}$$

et donc admet une unique mesure invariante. Ses transitions n'affectent qu'une station à la fois. Pour une station i donnée, les transitions possibles correspondent à deux évènements : le départ ou le retour d'un vélo.

— Départ d'un vélo : à une station i et au taux λ, un vélo disponible est remplacé par une place libre à un instant t. Ainsi $X_i^N(t) = \left(X_i^N(t^-) - 1\right) \mathbf{1}_{(X_i^N(t^-)>0)}$. Cet évènement a lieu si au

moins un vélo est disponible dans la station i. Dans le cas contraire, la transition n'a pas lieu et l'usager quitte le système.

— Retour d'un vélo : à la fin d'un trajet dont la durée est de loi exponentielle de paramètre μ, l'usager arrive à une station donnée i, choisie uniformément au hasard, à un instant t. Si une place est libre, le vélo est déposé à la station i. Une place libre se transforme alors en vélo disponible et donc $X_i^N(t) = \left(X_i^N(t^-) + 1\right) \mathbf{1}_{(X_i^N(t^-)<c)}$. Dans le cas contraire, l'usager commence un second trajet de même loi vers une station choisie au hasard parmi les N stations. Et ainsi de suite jusqu'à arriver à déposer le vélo.

Équations d'évolution stochastiques

Les dynamiques du processus $\left(X^N(t)\right)$ peuvent être exprimées en terme d'intégrales stochastiques par rapport à différents processus de Poisson. Introduisons les notations suivantes :

— Un processus de Poisson sur $\mathbb{R}_+$ d'intensité ξ est noté $\mathcal{N}_\xi$. Une suite i.i.d. de tels processus est notée $(\mathcal{N}_{\xi,i},\, i \in \mathbb{N})$

— un processus de Poisson marqué (t_n, U_n), où $(t_n, n \in \mathbb{N})$ est un processus de Poisson sur $\mathbb{R}_+$ d'intensité ξ et $(U_n, n \in \mathbb{N})$ une suite de variables aléatoires i.i.d. suivant une loi uniforme sur $\{1, \ldots, N\}$, est noté $\mathcal{N}_\xi^{U,N}$. Ainsi, pour $1 \leq i \leq N$, $\mathcal{N}_\xi^{U,N}(., \{i\})$ est un processus de Poisson sur $\mathbb{R}_+$ d'intensité ξ/N.

Suivant ces notations, pour $1 \leq i \leq N$, les départs des vélos de la station i sont modélisés par un processus de Poisson $\mathcal{N}_{\lambda,i}$ sur $\mathbb{R}_+$ d'intensité λ. Rappelons qu'à l'instant t^-, il y a $M - \sum_{k=1}^{N} X_k^N(t^-)$ vélos en circulation. Pour $1 \leq j \leq M - \sum_{k=1}^{N} X_k^N(t^-)$, le dépôt du vélo j à l'instant t est donc modélisé par un processus de Poisson marqué $\mathcal{N}_{\mu,j}^{U,N}$. Le choix de la station destination, disons $1 \leq i \leq N$, étant uniforme parmi les N stations, le retour du vélo j à la station i est modélisé par un saut du processus de Poisson $\mathcal{N}_{\mu,j}^{U,N}(., \{i\})$ d'intensité μ/N.

Ainsi, pour $1 \leq i \leq N$, l'équation différentielle stochastique régissant $X_i^N(t)$ s'écrit

$$dX_i^N(t) = -\mathbf{1}_{\{X_i^N(t^-)>0\}} \mathcal{N}_{\lambda,i}(dt) + \sum_{j=1}^{+\infty} \mathbf{1}_{\{X_i^N(t^-)<c\}} \mathbf{1}_{\{j \leqslant M - \sum_{k=1}^{N} X_k^N(t^-)\}} \mathcal{N}_{\mu,j}^{U,N}(dt, \{i\}).$$

En intégrant et en compensant les différents processus de Poisson, on obtient que

$$X_i^N(t) = X_i^N(0) - \lambda \int_0^t \mathbf{1}_{\{X_i^N(s)>0\}} ds + \frac{\mu}{N} \int_0^t \sum_{j=1}^{+\infty} \mathbf{1}_{\{X_i^N(s)<c\}} \mathbf{1}_{\{j \leqslant M - \sum_{k=1}^{N} X_k^N(s)\}} ds + M_i^N(t) \tag{1.1}$$

où le terme martingale $(M_i^N(t))$ est explicite.

Convergence Champ-moyen

Considérons le processus de mesure empirique sous sa forme "fonctionnelle" $\left(\Lambda^N(t)(f)\right)_t$ avec

$$\Lambda^N(t)(f) = \frac{1}{N} \sum_{i=1}^{N} f(X_i^N(t))$$

où f est une fonction à support fini dans $\mathbb{N}$. Partant de l'équation (1.1), on obtient que

$$df(X_i^N(t)) = (f(X_i^N(t^-)-1) - f(X_i^N(t^-)))\mathbf{1}_{\{X_i^N(t^-)>0\}}\mathcal{N}_{\lambda,i}(dt) + \sum_{j=1}^{+\infty}\mathbf{1}_{\{X_i^N(t^-)<c\}}\mathbf{1}_{\{j\leqslant M-\sum_{k=1}^N X_k^N(t^-)\}}(f(X_i^N(t^-)+1)-f(X_i^N(t^-)))\mathcal{N}_{\mu,j}^{U,N}(dt,\{i\}).$$

On notera dans la suite $\Delta^{\pm}f(x) = f(x\pm 1) - f(x)$ de sorte qu'on obtient que

$$f(X_i^N(t)) = f(X_i^N(0)) + \lambda\int_0^t \Delta^- f(X_i^N(s))\mathbf{1}_{\{X_i^N(s)>0\}}ds + \frac{\mu}{N}\int_0^t\sum_{j=1}^{+\infty}\Delta^+ f(X_i^N(s))\mathbf{1}_{\{X_i^N(s)<c\}}\mathbf{1}_{\{j\leqslant M-\sum_{k=1}^N X_k^N(s)\}}ds + \mathcal{M}_{i,f}^N(t) \tag{1.2}$$

où le terme martingale est donné par

$$\mathcal{M}_{i,f}^N(t) = \int_0^t \Delta^- f(X_i^N(s))\mathbf{1}_{\{X_i^N(s)>0\}}\left(\mathcal{N}_{\lambda,i}(ds)-\lambda ds\right) + \int_0^t\sum_{j=1}^{+\infty}\Delta^+ f(X_i^N(s))\mathbf{1}_{\{X_i^N(s)<c\}}\mathbf{1}_{\{j\leqslant M-\sum_{k=1}^N X_k^N(s)\}}\left(\mathcal{N}_{\mu,j}^{U,N}(ds,\{i\})-\frac{\mu}{N}ds\right).$$

En moyennant sur N l'équation (1.2) et en notant la martingale moyenne $\mathcal{M}_f^N(t)$, on obtient l'équation d'évolution de $\Lambda^N(t)(f)$ écrite sous la forme compacte suivante

$$\Lambda^N(t)(f) = \Lambda^N(0)(f) + \lambda\int_0^t \Lambda^N(s)(\Delta^- f\mathbf{1}_{\{\mathbb{N}^*\}})ds + \mu\int_0^t\left(\frac{M}{N}-\Lambda^N(s)(Id_{[0,c]})\right)\Lambda^N(s)(\Delta^+ f\mathbf{1}_{\{0,\dots,c-1\}})ds + \mathcal{M}_f^N(t) \tag{1.3}$$

où $Id_{[0,c]}$ est la fonction identité restreinte à $[0,c]$.

Convergence champ-moyen vers un processus déterministe

Soit $T>0$ fixé. Pour établir la convergence champ-moyen de la suite des processus de mesure empirique $(\Lambda^N(t))_{0\leqslant t\leqslant T}$, la technique est assez standard (voir [EK86]) : tension de la suite $(\Lambda^N(t))$ et unicité de la valeur d'adhérence.
On montre d'abord que la suite des mesures empiriques est tendue pour la topologie de Skorokhod en utilisant le critère du module de continuité (voir [Rob13] par exemple) où il suffit de montrer que pour tous $\varepsilon>0$ et $\eta>0$, il existe $\delta_0>0$ et un rang $N_0\in\mathbb{N}$ tels que pour tous $\delta<\delta_0$ et $N\geqslant N_0$, on a

$$\mathbb{P}\left(\sup_{\substack{0\leqslant s,t\leqslant T\\ |t-s|\leqslant\delta}}\left|\Lambda^N(t)(f)-\Lambda^N(s)(f)\right|\geqslant\eta\right)<\epsilon.$$

Soient $s,t\in[0,T]$ tels que $|s-t|<\delta$ pour un $\delta>0$ à déterminer. Il est simple de borner les deux intégrales apparaissant dans (1.3). En effet, en utilisant le fait que le terme $\left(M/N-\Lambda^N(s)(Id_{[0,c]})\right)$ soit borné puisque $M/N\sim s$, on obtient assez facilement l'existence d'une constante $K>0$ telle que

$$\left|\Lambda^N(t)(f)-\Lambda^N(s)(f)\right| \leq \delta K\|f\| + \left|\mathcal{M}_f^N(t)-\mathcal{M}_f^N(s)\right|.$$

Le calcul direct du processus croissant du terme martingale dans (1.3) permet d'établir que

$$\lim_{N\to+\infty} \mathbb{E}\left[\langle \mathcal{M}_f^N\rangle(T)\right] = 0.$$

En utilisant les inégalités de Cauchy-Schwarz et de Doob, on arrive à

$$\mathbb{E}\Big(\sup_{\substack{0\leqslant s,t\leqslant T\\ |t-s|\leqslant\delta}} \left|\Lambda^N(t)(f) - \Lambda^N(s)(f)\right|\Big) \leq \delta K\|f\| + O(1/N).$$

Il est alors possible de choisir δ_0 et N_0 de telle sorte que le terme de droite de cette inégalité soit inférieur au seuil ε fixé pour tous $\delta < \delta_0$ et $N \geqslant N_0$. On obtient ainsi la propriété de tension voulue. Le terme martingale dans (1.3) disparaît lorsque N devient grand impliquant que toute valeur d'adhérence du processus $(\Lambda^N(t))$ vérifie nécessairement l'équation suivante :

$$\begin{aligned}\Lambda(t)(f) = \Lambda(0)(f) &+ \lambda\int_0^t \Lambda(s)(\Delta^- f\mathbf{1}_{\{\mathbb{N}^*\}})ds \\ &+ \mu\int_0^t \Big(s - \Lambda(s)(Id_{[0,c]})\Big)\Lambda(s)(\Delta^+ f\mathbf{1}_{\{0,\dots,c-1\}})ds\end{aligned} \tag{1.4}$$

pour toute fonction test f à support fini dans $\mathbb{N}$.
En prouvant que l'équation (1.4) admet au plus une solution, on obtient la convergence en distribution de la suite $(\Lambda^N(t))$. Un raisonnement par l'absurde ajouté au lemme de Gronwall suffit pour conclure à cette unicité. On obtient alors ce premier résultat.

Théorème 1.2.1 (*Convergence champ-moyen*)

Pour $T > 0$, la suite des processus des mesures $(\Lambda^N(t)(f))_{0\leqslant t\leqslant T}$ converge en distribution vers un processus déterministe $(\Lambda(t)(f))_{0\leqslant t\leqslant T}$ unique solution de l'équation (1.4) *pour toute fonction f à support fini dans $\mathbb{N}$.*

Une interprétation files d'attente

La proportion de stations à k vélos à un instant $t \geqslant 0$ avec $k \in \{0, ..., c\}$ est définie par

$$Y_k^N(t) := \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{(X_i^N(t)=k)}.$$

Considérons le processus empirique associé $(Y^N(t)) := (Y_0^N(t), ..., Y_c^N(t))_{t\geqslant 0}$. Il peut être réécrit comme suit

$$Y^N(t) = (Y_0^N(t), ..., Y_c^N(t)) = \sum_{k=0}^c Y_k^N(t)e_k = \frac{1}{N}\sum_{i=1}^N\sum_{k=0}^c \mathbf{1}_k(X_i^N(t))e_k = \Lambda^N(t)(f)$$

où $(e_k)_{0\leq k\leq c}$ désigne la base canonique de $\mathbb{R}^{c+1}$ et $f(x) = \sum_{k=0}^c \mathbf{1}_k(x)\, e_k$. L'équation d'évolution de $(Y^N(t))$ est alors un cas particulier de (1.3). D'après le Théorème 1.2.1, on sait que le processus $(Y^N(t))$ converge en distribution vers un certain processus déterministe $(y(t))$, unique solution de l'équation limite (1.4) pour notre fonction test f. Ainsi, par passage à la limite quand N devient grand dans (1.3), où on utilise le fait que

$$\Lambda^N(s)(Id_{[0,c]}) = \frac{1}{N}\sum_{i=1}^N X_i(s) = \sum_{k=0}^c kY_k(s),$$

on obtient que

$$y(t) = y(0) + \int_0^t \sum_{k=0}^{c} y_k(s)\Big(\lambda \mathbf{1}_{\{k>0\}}(e_{k-1} - e_k) + \mu(s - m(y(s)))\mathbf{1}_{\{k<c\}}(e_{k+1} - e_k)\Big)ds \quad (1.5)$$

où on introduit la notation suivante $m(y(s)) := \sum_{k=0}^{c} k y_k(s)$ pour tout $s \geqslant 0$. L'équation limite (1.5) vérifiée par $(y(t))$ est bien similaire à celle énoncée dans [FG16]. C'est un système dynamique sur l'ensemble des probabilités sur $\{0, \dots, c\}$ pouvant être réécrit d'une manière plus compacte :

$$y'(t) = y(t)\, L_{y(t)}$$

où $L_y(t)$ est le générateur infinitésimal d'une file $M/M/1/c$ dont le taux d'arrivée est $\mu(s - m(y(t)))$ (taux d'arrivée à une station donnée d'un vélo qui roule) et le taux de service λ (taux d'arrivée des usagers à une station donnée).
L'approximation champ-moyen veut dire que, quand le système devient grand, une station donnée se comporte à un instant $t \geqslant 0$ comme une file à un serveur et à capacité finie c où les clients (les vélos) arrivent à la file (la station) au taux $\mu(s - m(y(t)))$ et la quittent au taux λ. Cette interprétation files d'attente est précieuse puisque dans le cas général d'un réseau homogène à N noeuds (sites, stations, . . .) de capacité finie c, le système dynamique obtenu est de la forme

$$y'(t) = \mathcal{V}(y(t))$$

où $\mathcal{V}$ est un champ de vecteur sur l'ensemble des probabilités $\mathcal{P}_{\{0,\dots,c\}}$. L'écriture particulière de ce champ de vecteurs comme $\mathcal{V}(y) = y\, L_y$ où L_y est le générateur infinitésimal d'une file $M/M/1/c$ pour notre modèle implique qu'un point d'équilibre $\overline{y}$ du système dynamique, c'est-à-dire solution de $y\, L_y = 0$ est la mesure invariante associée à L_y. C'est donc une loi géométrique tronquée de paramètre le rapport entre le taux d'arrivée et celui de départ de la file typique associée au générateur $L_{\overline{y}}$, soit $\pi_{\rho,c} \sim Geom(\rho(y))$ sur $\{0, \dots, c\}$ avec

$$\rho(y) = \frac{\mu(s - m(y))}{\lambda}. \quad (1.6)$$

et $m(y) = \sum_{k=0}^{c} k y_k$. Ainsi, pour tout $k \in \{0, \dots, c\}$, on a $\pi_{\rho,c}(k) = \dfrac{1-\rho}{1-\rho^{c+1}}\rho^k$ où ρ par définition est solution de l'équation de point fixe (1.6) réécrite aussi

$$s = \frac{\lambda}{\mu}\rho + \sum_{k=0}^{c} k\, \pi_{\rho,c}(k). \quad (1.7)$$

L'unicité du point fixe ρ sur $]0, +\infty[$ est due à un argument de monotonie.

Remarque 1.2.1

- *Le cas $\rho = 1$ correspond à une loi uniforme sur $\{0, \dots, c\}$. Dans ce cas, pour tout $0 \leq k \leq c$, on a $\pi_{\rho,c}(k) = 1/(c+1)$.*
- *Remarquons que cette interprétation files d'attente nous permet de passer d'une équation de point fixe vectorielle $y\, L_y = 0$ avec $y \in \mathbb{R}^{c+1}$ vérifiée par la mesure invariante à une équation de point fixe réelle sur le paramètre de cette loi invariante.*

Convergence vers le point d'équilibre

L'existence et l'unicité du point d'équilibre du système dynamique limite (obtenu quand N tend vers l'infini) ne suffit pas pour obtenir la convergence de la mesure invariante du processus mesure empirique $(Y^N(t))$ vers ce point. Dans l'idéal, ce résultat est obtenu grâce à une fonction de Lyapunov qui garantit la convergence de toutes les trajectoires solutions du système dynamique limite vers le point fixe. Un moyen pour construire une telle fonction est de considérer un terme d'entropie relative plus un terme correctif comme dans [Gas16]. Exhiber une fonction de Lyapunov n'est pas un problème facile. Dans [FGM12], on propose une fonction de Lyapunov dans un cadre hétérogène.

Retour à l'application : le problème de dimensionnement

Rappelons que le paramètre $s := \lim_{N\to+\infty} M/N$ est un paramètre clé pour l'opérateur puisqu'il représente le nombre de vélos par station, en moyenne. L'équation (1.7) dit que ce nombre se répartit entre une partie de vélos en trajet $(\rho\lambda/\mu)$ et d'une autre partie en station $(\sum_{k=0}^{c} k\,\pi_\rho(k))$. Si on se fixe comme métrique la proportion limite de stations problématiques, c'est-à-dire vides (sans vélos) ou saturées (sans place de parking), ceci revient à considérer $Pb(\rho) = \pi_\rho(0) + \pi_\rho(c)$. L'étude de la courbe paramétrique $\rho \mapsto (s(\rho), Pb(\rho))$ exprimant cette proportion de stations problématiques en fonction du ratio de dimensionnement s permet d'obtenir la taille de la flotte maximisant la performance du système, autrement dit le ratio s^* minimisant la proportion de stations problématiques. On retrouve ainsi le résultat établi dans [FG16]

$$s^* = \frac{c}{2} + \frac{\lambda}{\mu}.$$

Ceci revient à dire que l'opérateur aurait intérêt à "remplir" les stations à moitié avec un paramètre d'ajustement qui dépend à la fois du taux d'arrivée des usagers et du temps de trajet moyen. Plus il y a de demande ou plus les vélos roulent, plus il a intérêt à augmenter la taille de sa flotte. Cet optimum correspondant à $\rho = 1$, soit une mesure invariante uniforme sur $\{0, \dots, c\}$. Intuitivement, ce résultat semble naturel puisque le modèle étudié est homogène.

1.3 Le cas hétérogène

Considérons un système de vélos en libre-service avec N stations et une flotte de M vélos. Chaque station i a une capacité $c_{i,N}$. La dynamique du système est la suivante. Les usagers arrivent aux stations selon des processus de Poisson indépendants avec un taux $\lambda_{i,N}$ à la station i. Lorsqu'un usager arrive à une station donnée sans vélo disponible, il quitte le système. Dans le cas contraire, il prend un vélo et choisit la station j avec la probabilité $p_{j,N}$. La durée de trajet a une distribution exponentielle de paramètre μ_N, quelle que soit la station d'où il vient. Lorsqu'il arrive à la station j, s'il y a moins de $c_{j,N}$ vélos dans cette station, il dépose son vélo et quitte le système. S'il y a $c_{j,N}$ vélos (c'est-à-dire que la station est saturée), l'usager choisit à nouveau une station, disons k, avec une probabilité $p_{k,N}$ et se rend à cette station. La durée de ce second trajet est de distribution exponentielle avec le même paramètre μ_N. Et ainsi de suite jusqu'à ce qu'il puisse rendre son vélo.
Ce modèle est une généralisation du cas homogène où les taux d'arrivée ne dépendent pas des stations et où l'usager rend le vélo à une station choisie uniformément au hasard. Même si les différentes probabilités d'aller d'une station à l'autre ne constituent pas vraiment une matrice de routage, la popularité d'une station est prise en compte.
Pour N fixé, notons par $R_{i,N} := \mu_N p_{i,N}/\lambda_{i,N}$ un paramètre caractéristique de la station i appelée *utilisation*. Ce paramètre apparaît dans les taux des transitions du processus d'état du système $\left(X^N(t)\right) := \left(X_i^N(t), 1 \leqslant i \leqslant N\right)$ où $X_i^N(t)$ est le nombre de vélos dans la station i à l'instant t.

Le processus de Markov $\left(X^N(t)\right)$ est irréductible sur un espace d'états fini. L'approche champ-moyen exhibée dans le cas homogène s'applique au cas hétérogène, sous une condition supplémentaire.

Hypothèse 1.3.1

Il existe une mesure de probabilités I sur $]0;1] \times \mathbb{N}$ et une constante $\gamma > 0$ telles que, quand N tend vers l'infini,

$$\frac{1}{N}\sum_{i=1}^{N} \delta_{(r_{i,N};c_{i,N})} \xrightarrow{(w)} I$$
$$N\, R_{max,N} \to 1/\gamma$$

où $R_{max,N} := \max_i R_{i,N}$ et $r_{i,N} := R_{i,N}/R_{max,N}$ désigne l'utilisation relative de la station i. La convergence (w) s'entend au sens de la convergence faible.

Notons que cette hypothèse est liée à la topologie du système (voir les exemples ci-dessous). Sous l'hypothèse précédente, on obtient dans [FGM12] le résultat suivant

Théorème 1.3.1 (*Convergence champ-moyen*)

Sous l'hypothèse 1.3.1, si I est à support fini, alors lorsque N et M tendent vers l'infini avec M/N tendant vers un certain s, le nombre de vélos à une station avec les paramètres r (utilisation relative) et c (capacité) a une distribution stationnaire géométrique $\pi_{\rho r,c}$ sur $\{0,...,c\}$ avec le paramètre ρr où ρ est l'unique solution de

$$s = \rho\gamma + \int_{]0;1]\times\mathbb{N}} m(\pi_{\rho r,c}) dI(r,c) \tag{1.8}$$

où la moyenne d'une variable aléatoire avec une distribution π est notée $m(\pi)$.

Pour simplifier ce résultat, prenons deux cas.

— Le cas homogène : ceci implique que pour tout i, $c_{i,N} = c$, $\lambda_{i,N} = \lambda$, $p_{i,N} = 1/N$ et $\mu_N = \mu$ de telle façon que $r_{i,N} = 1$. L'hypothèse 1.3.1 devient triviale puisque $I = \delta_{1,c}$ et $\gamma = \lambda/\mu$. L'équation (1.8) devient alors

$$s = \rho\lambda/\mu + m(\pi_{\rho,c})$$

ce qui coïncide avec l'équation (1.7).

— Le cas à K clusters : Supposons que les N stations forment K clusters tel que chaque cluster $1 \le k \le K$ a N_k stations avec N_k/Nqui tend vers α_k. Les stations du cluster k ont toutes les mêmes paramètres λ_k, c_k et $p_{k,N} = \beta_k/N$. Ainsi, l'hypothèse 1.3.1 est vraie avec $I = \sum_{k=1}^{c} \alpha_k \delta_{(r_k;c_k)}$, $r_k = \gamma\mu\beta_k/\lambda_k$ et $\gamma = 1/\max_k(\mu\beta_k/\lambda_k)$. L'équation (1.8) devient alors

$$s = \rho\gamma + \sum_{k=1}^{c} \alpha_k m(\pi_{\rho r_k,c_k}).$$

Cet exemple modélise une zone de service d'un système de vélopartage où le découpage en clusters est fait sur la base de la demande (forte demande, faible demande) alors que le temps de trajet moyen est constant.

1.4 Une flotte mixte : vélos mécanique et vélos électriques

Les vélos électriques se sont déployés massivement dans les systèmes de vélos en libre-service préexistants afin d'attirer de nouveaux usagers et de remplacer les voitures à plus grande échelle.

Mais cela entraîne des interactions entre les deux populations de vélos. Dans [AFM22], on propose un modèle de système homogène de partage de vélos où les deux classes de vélos interagissent uniquement à travers le partage de la capacité finie des stations. Il modélise les systèmes avec des vélos électriques et mécaniques nécessitant des abonnements différents. Il s'agit d'une première analyse stochastique à grande échelle pour des systèmes de vélopartage à flotte mixte. L'approche champ-moyen nous permet d'établir explicitement la distribution limite de l'état d'une station lorsque le nombre de stations N et la taille de la flotte de chaque classe M_1 et M_2 augmentent au même rythme, c'est-à-dire $M_l/N \sim s_l$ pour $l \in \{1, 2\}$.
Si on considère, dans un cadre homogène, le processus $(X^N(t) := X^N_{l,i}(t),\, 1 \leq i \leq N,\, l = 1, 2)$ où $X^N_{l,i}(t)$ désigne le nombre de vélos de type $l \in \{1, 2\}$ à l'instant t dans la station i, alors on peut montrer d'une manière similaire au modèle homogène de vélopartage à une seule classe que le processus empirique $(Y_{n_1,n_2}(t))_{t\in[0,T]}$ défini, pour $n_1 + n_2 \leq c$, par

$$Y_{n_1,n_2} := \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\{X^N_{1,i}(t)=n_1, X^N_{2,i}(t)=n_2\}}$$

converge en distribution vers une fonction déterministe $(y(t))_{t\in[0,T]}$, solution d'une EDO explicite. Le point d'équilibre associé à ce système dynamique est identifié à la mesure invariante d'une file d'attente $M/M/1/c$ à deux classes de clients puisque l'interaction entre les deux types de clients (vélos) se limite au partage de la capacité globale c (de la station). Ainsi, on obtient l'existence et l'unicité de la mesure stationnaire, une distribution géométrique bivariée tronquée

$$\pi_{\rho_1,\rho_2}(n_1, n_2) = \frac{\rho_1^{n_1}\rho_2^{n_2}}{Z(\rho_1, \rho_2)}$$

où les paramètres ρ_1 et ρ_2 sont déterminés de manière unique lorsque les proportions limites s_1 et s_2 de chaque type de vélos ainsi que la capacité c de chaque station sont fixées.
Pour répondre à la question du nombre optimal de vélos de chaque type par station pour une capacité c donnée, il faudra aussi préciser le critère de performance considéré comme étant la proportion de stations *problématiques* Pb, c'est-à dire des stations saturées ou ne comportant aucun vélo d'un des deux types. Néanmoins, on se heurte à une relation implicite reliant les deux paramètres ρ_1 et ρ_2, rendant l'obtention d'une formule explicite des paramètres de flotte optimaux s_1^* et s_2^* hors de portée. Le cas symétrique où $\lambda_1 = \lambda_2$, $\mu_1 = \mu_2$ et $s_1 = s_2 = s/2$ implique que $\rho_1 = \rho_2 = \rho$. Il est alors possible d'exprimer explicitement s et Pb en fonction de ρ. Ce cas a été traité dans [AFM22] où on prouve le résultat intuitif que le ratio optimal pour chaque classe correspond au tiers de la capacité (un tiers pour chaque ressource : vélos mécaniques, vélos électriques et places de parking) plus un terme correctif qui dépend du trafic, à l'instar du modèle homogène avec un seul type de vélos.

Théorème 1.4.1 (*Performance - Le cas symétrique*)

La proportion minimale de stations problématiques pour ce modèle dans le cas symétrique est $P_b^* = 6c/((c+1)(c+2))$ *et est atteinte quand* $s^* = 2\lambda/\mu + 2c/3$. *Le* ρ *correspondant est* 1.

En dehors de ce cas totalement symétrique, la relation entre ρ_1 et ρ_2 est implicite. Les méthodes de résolution numériques permettent d'obtenir les paramètres ρ_1 et ρ_2 et donc la proportion minimale de stations problématiques ainsi que les paramètres de flottes qui lui correspondent.

Un modèle plus réaliste : changement de classe

Limiter l'interaction entre les deux classes de clients (les deux types de vélos) au partage de la capacité globale est loin de la réalité. En cas de manque de l'une des deux ressources, il est souhaitable de

considérer qu'une proportion α_i des usagers d'un type de vélos se reporte sur l'autre type. L'écriture des transitions du même processus de mesure empirique que initialement ainsi que la convergence champ-moyen restent valide. Néanmoins, on perd l'interprétation files d'attente qui permettait la caractérisation simple du point d'équilibre $\overline{y}$. Une tentative timide d'une approche analytique via la méthode du noyau reste inachevée. A la manière de [FIM17], on considère la fonction génératrice $F(x,y) = \sum_{k+l\leqslant c} \overline{y}_{k,l} x^k y^l$. Cette fonction vérifie une équation fonctionnelle, à laquelle on ajoute la condition de normalisation $F(1,1) = 1$ pour caractériser F et donc $\overline{y}$. Cependant, comme les conditions aux bords $F(.,0)$, $F(0,.)$ et $F_c : (x,y) \mapsto \sum_{k+l=c} \overline{y}_{k,l} x^k y^l$ sont inconnues, déterminer F reste hors de portée. Cette démarche est en dehors du cadre théorique abordé dans ce manuscrit.

1.5 Une politique incitative

Je présente ici brièvement un modèle stochastique visant à étudier une politique incitative que l'opérateur peut mettre en place afin de prévenir le manque en vélos. En effet, en offrant par exemple des minutes gratuites aux usagers qui ramènent, vers une zone de forte demande telle que le centre ville par exemple, des vélos inactifs depuis un certain temps, garés en périphérie, l'opérateur espère générer une proportion de "bons" trajets afin de réguler le système et diminuer son déséquilibre. Ces vélos "isolés" peuvent être mis en avant par l'opérateur et désignés sur son application comme des cadeaux. Cette politique cadeaux est mise en place par Communauto pour son système d'autopartage Flex à Montréal où l'opérateur offre 30 minutes gratuites aux usagers qui ramènent des voitures qu'il désigne comme cadeaux vers une zone de forte demande. L'impact de la politique cadeaux de Communauto Montréal sur le comportement des usagers est discuté dans [MFM+24] à partir d'une analyse de données opérateur. Néanmoins, la très faible proportion de trajets cadeaux proposés par l'opérateur et de cadeaux réellement utilisés par les usagers dans le volume total des transactions rend difficile l'observation de son impact. L'intérêt d'une modélisation stochastique du système pour établir des résultas qualitatifs et quantitatifs de cette politique incitative est réel pour l'opérateur. Pour des raisons de cohérence, je préfère aborder cette politique incitative ici puisqu'elle peut être appliquée à tout système de partage de véhicules (vélos, scooters, trottinettes, voitures ...).

Le modèle

Nous optons alors pour un cadre homogène avec deux clusters où les paramètres des stations sont les mêmes pour chaque cluster. Les dynamiques du modèle sont les suivantes.

- Un usager arrive à une station du cluster $i \in \{1,2\}$ avec un taux λ_i. La zone de forte demande est désignée par cluster 1, celle normale par cluster 2. Comme le taux d'arrivée des usagers est plus important dans le premier cluster, alors $\lambda_1 > \lambda_2$.
- Si l'usager arrive à une station du cluster 1 où il y a un véhicule disponible, il le prend pour un trajet. Sinon, il quitte le système.
- Chaque véhicule garé dans le cluster 2 a une horloge. Lorsque cette horloge sonne pour un véhicule, celui-ci devient un *cadeau*.
- Lorsqu'un usager arrive à une station du cluster 2, s'il y a un *cadeau* disponible et une véhicule normal (non cadeau) disponible dans cette station, il prend le véhicule cadeau avec une probabilité p, et le véhicule normal avec une probabilité $1-p$. S'il n'y a qu'une seule ressource (cadeau ou non), l'usager prend ce qui est disponible. Sinon, il quitte le système.
- À la fin du trajet, l'usager utilisant une véhicule normal choisit le cluster 1 avec une probabilité r , respectivement le cluster 2 avec une probabilité $1-r$, puis il choisit une station au hasard dans ce cluster pour garer le véhicule.

– Lorsqu'un trajet-cadeau se termine, l'usager rend la voiture-cadeau à n'importe quelle station du cluster 1 avec une probabilité q, respectivement du cluster 2 avec une probabilité $1-q$. Le véhicule-cadeau garé apparaît alors comme un véhicule normal sur l'appli.
– Une station du cluster i a une capacité finie c_i. Si la station choisie est saturée, l'usager fait un autre trajet jusqu'à trouver une station avec une place de parking disponible.

La Figure 1.1 illustre les dynamiques du modèle.

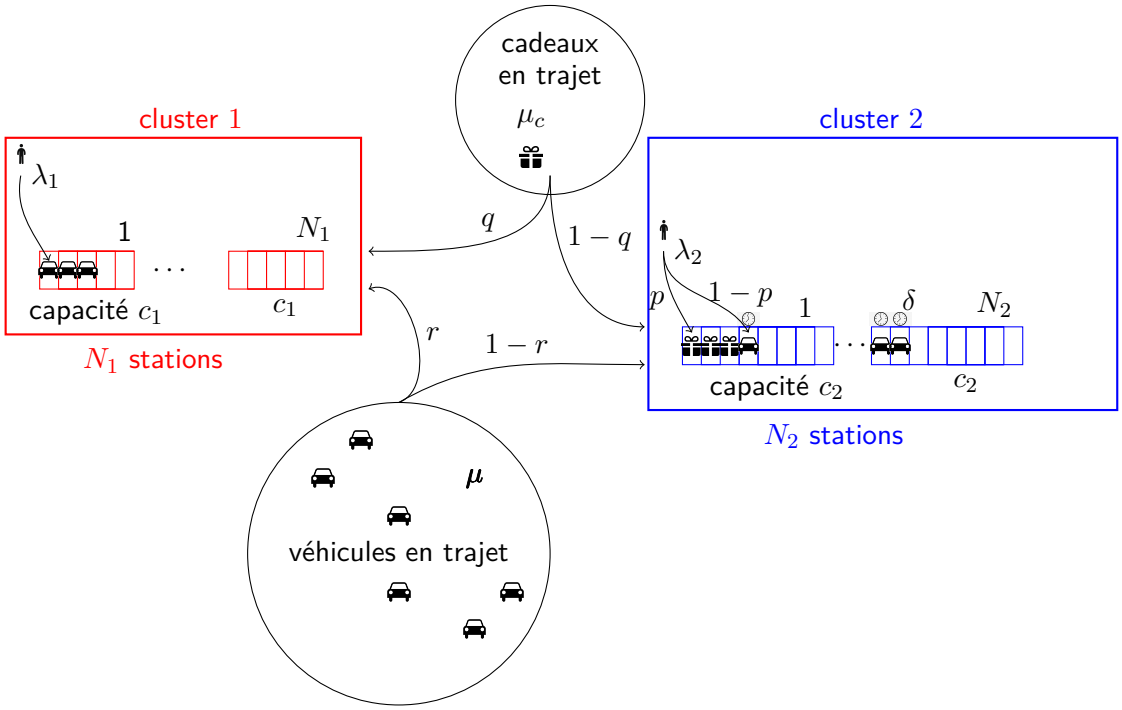

Figure 1.1 – Illustration des dynamiques du modèle avec cadeaux.

Voici les notations du modèle. L'indice $i \in \{1, 2\}$ désigne le type du cluster (zone).

– N_i est le nombre de stations dans le cluster i.
– $N = \sum_i N_i$ est le nombre totale de stations.
– $\alpha_i = N_i/N$ est la proportion des stations du cluster i.
– c_i est la capacité d'une station du cluster i.
– M est le nombre total de véhicules partagés.
– $s = M/N$ est le nombre moyen de véhicules par station.
– λ_i est la taux d'arrivée des usagers à une station du cluster i.
– $1/\mu$ est le temps de trajet moyen d'un véhicule normal (non cadeau).
– $1/\mu_c$ est le temps de trajet moyen d'un véhicule *cadeau*.
– δ est le taux auquel un véhicule normal garé dans une station du cluster 2 devient *cadeau*.
– p est la probabilité qu'un usager choisisse un véhicule cadeau.
– q est la probabilité que l'usager ramène le véhicule cadeau vers une station du cluster 1.
– r est la probabilité qu'un usager ramène un véhicule normal vers une station du cluster 1.

Le système peut être décrit comme un réseau stochastique fermé. Les nœuds du réseau sont un ensemble de $N = N_1 + N_2$ files d'attente $M/M/1/c_i$ (où $i = 1, 2$) à capacité finie, les stations, divisées en deux clusters, le cluster 1 (pour la zone à haute demande) avec N_1 stations de capacité c_1, le cluster 2 (pour la zone normale) avec N_2 stations de capacité c_2, plus deux files d'attente

$M/M/\infty$ représentant respectivement les véhicules normaux et ceux cadeaux en route. Les temps de service aux files d'attente ont une distribution exponentielle avec des paramètres respectifs λ_1, λ_2, μ et μ_c. Selon le vocabulaire des files d'attente, il y a M clients de deux classes : véhicules et cadeaux, et une matrice de routage donnée par la description précédente. Cependant, il ne s'agit pas d'un réseau de Jackson car il y a des transitions supplémentaires impliquant un changement de classe : un véhicule dans un nœud $M/M/1/c_2$ (une station) du cluster 2 devient un cadeau au taux δ et un cadeau arrivant à une station depuis le nœud $M/M/\infty$ (la route) devient un véhicule normal.

Le processus d'état

Le processus d'état est noté par

$$(X_{1,i}(t), X_{2,j}(t), C_j(t), R^N(t),\ 1 \leqslant i \leqslant N_1 \ \text{ et } \ 1 \leqslant j \leqslant N_2)$$

où

- $X_{1,i}(t)$ est le nombre de véhicule à une station i du cluster 1 à l'instant t,
- $X_{2,j}(t)$ est le nombre de véhicule à une station j du cluster 2 à l'instant t,
- $C_j(t)$ est le nombre de cadeaux à la station j (nécessairement du cluster 2) à l'instant t,
- $R^N(t)$ est le nombre de cadeaux en trajet à l'instant t.

Le nombre de véhicules en trajet à l'instant t est égal à $M - \sum_{i=1}^{N_1} X_{1,i}(t) - \sum_{j=1}^{N_2} X_{2,j}(t) - R^N(t)$. Comme le modèle est homogène, considérons le processus de mesure empirique

$$(Y^N(t)) = \left(Y^{N_1}_{1,k}(t), Y^{N_2}_{2,k',l'}(t), \frac{R^N(t)}{N}, k \in \chi_1, (k',l') \in \chi_2\right)$$

où $Y^{N_1}_{1,k}(t)$ est la proportion de stations à k véhicules dans le cluster 1 et $Y^{N_2}_{2,k',l'}(t)$ est la proportion de stations à k' véhicules et l' cadeaux dans le cluster 2, définis par

$$Y^{N_1}_{1,k}(t) = \frac{1}{N_1}\sum_{i=1}^{N_1} \mathbf{1}_{\{X_{1,i}(t)=k\}} \text{ et } Y^{N_2}_{2,k',l'}(t) = \frac{1}{N_2}\sum_{i=1}^{N_2} \mathbf{1}_{\left\{\left(X_{2,i}(t),C_i(t)\right)=(k',l')\right\}}$$

avec

$$\chi_1 = \{k \in \mathbb{N},\ k \leqslant c_1\}, \quad \chi_2 = \{(k,l) \in \mathbb{N}^2,\ k+l \leqslant c_2\}.$$

Toutes les durées de temps suivent des lois exponentielles indépendantes, on a alors que $(Y^N(t))$ est un processus de Markov sur un espace d'états fini. On obtient, d'une manière similaire à celle exposée pour un modèle homogène de vélopartage à une classe de clients, que lorsque N tend vers l'infini, le processus $(Y^N(t))$ converge en distribution vers une fonction déterministe, unique solution d'un système explicite d'équations différentielles ordinaires de la forme

$$\frac{dy}{dt}(t) = F(y(t)).$$

Déterminer le point d'équilibre revient à résoudre un système non linéaire $F(\overline{y}) = 0$ qui semble hors de portée. Il existe de nombreux outils pour résoudre numériquement son point d'équilibre. Dans [MFM+22], nous avons opté pour la méthode d'Anderson implémentée dans Scipy, une bibliothèque Python, pour une résolution numérique qui nous a permis de discuter l'influence des paramètres du modèle, notamment la durée au bout de laquelle le véhicule est désigné cadeau par l'opérateur (le paramètre δ), ou encore l'impact de l'adhésion d'une proportion donnée des usagers à la politique incitative (les paramètres p et q).

1.6 Et l'aspect local ?

En cas de saturation de la station d'arrivée, nous avons toujours considéré que l'usager recommence un autre trajet vers une station choisie uniformément au hasard parmi les N stations. Cette hypothèse, bien que peu réaliste, n'empêche pas l'obtention d'un comportement théorique très similaire à celui obtenu par simulation des vraies dynamiques observées sur les données opérateur. Dans [DFM18], nous avons voulu analyser un modèle qui tient compte du choix local quand l'usager choisit la station la moins remplie parmi un petit voisinage de sa destination. Ainsi, [DFM18] traite de l'impact du choix entre deux voisins dans un grand ensemble de files d'attente. Le choix est une politique d'équilibrage des charges parmi d'autres telles que *offloading*, *redundancy* ou encore *work stealing* ([GLM+10], [GHBSW+17], [GB10] et autres) par exemple. La politique du choix parmi deux est une méthode distribuée bien connue pour son efficacité. Pour cette politique, les clients qui arrivent choisissent deux files d'attente au hasard et rejoignent la plus courte, les égalités étant résolues de manière aléatoire. Voir [VDK96] et [Mit96] pour un réseau de files d'attente à un serveur.

Le modèle que je présente ici est appelé modèle de choix *local*. Il consiste en un ensemble de N files d'attente à un serveur avec une capacité infinie où les clients arrivent à chaque file selon des processus de Poisson indépendants avec un taux λ. Lorsqu'un client arrive à la file i, $1 \leq i \leq N$, il choisit entre les files i et $i+1$ celle qui est la moins chargée et s'y inscrit. Par convention, la file $N+1$ est la file 1. Si les files i et $i+1$ ont le même nombre de clients, il rejoint l'une des deux avec une probabilité $1/2$. Les temps de service sont i.i.d. avec une distribution exponentielle de paramètre μ. Lorsque le client est servi, il quitte le système. Tous les temps d'inter-arrivée et de service sont indépendants. La charge ρ est par définition λ/μ. Notre intérêt concerne la distribution marginale du nombre de clients dans une file d'attente à l'équilibre pour le modèle de choix local quand le nombre N de files d'attente est fixé. Nous étudions l'asymptotique des probabilités stationnaires pour une file d'attente lorsque la charge tend vers zéro pour une comparaison avec le modèle de choix aléatoire, dans lequel un client arrivant choisit uniformément deux files au hasard parmi N et rejoint celle qui est la moins chargée, ainsi que le modèle sans choix, dans lequel un client arrivant dans la file i est servi dans cette file. Le modèle sans choix est simplement constitué de N files d'attente $M/M/1$ indépendantes dont la distribution stationnaire de la longueur de la file est géométrique de paramètre ρ pour $\rho < 1$. Pour le modèle de choix aléatoire, la limite, lorsque N devient grand, de la probabilité stationnaire qu'une file d'attente ait plus de k clients, est doublement exponentiellement, plus précisément est ρ^{2^k-1}, pour $k \geqslant 0$. Cette décroissance doublement exponentielle est connue dans la littérature sous le nom de puissance du choix étant considérablement plus petite que la probabilité de queue ρ^k, pour le modèle sans choix.

L'approche est différente puisqu'il ne s'agit plus de limite champ-moyen (faire tendre N vers l'infini) suivi de la recherche de point d'équilibre pour établir le comportement en temps long (t tend vers l'infini) du système. Dans notre cas ici, N est fixé et on note par $(X(t) := (X_i(t),\, 1 \leq i \leq N)$, où $(X_i(t))$ est le nombre de clients à la file d'attente i à l'instant t, le processus de longueur des files d'attente. $(X(t))$ est un processus de Markov sur l'espace d'états $\mathbb{N}^N$. Dans [DFM18], on prouve que la condition $\rho < 1$ garantit que le processus $(X(t))$ est ergodique. Soit $y = (y_n,\, n \in \mathbb{N}^N)$ sa mesure invariante, unique solution des équations de balance globale. L'argument majeur de notre approche est de considérer, pour tout $n \in \mathbb{N}^N$, la mesure de probabilité y_n comme fonction analytique du paramètre ρ sur un voisinage de zéro. Une procédure d'induction fournit alors tous les termes de la série entière. Cette hypothèse de l'analyticité des probabilités stationnaires d'une famille de chaînes de Markov dépendant d'un paramètre est la principale question abordée par [MM79] (voir également [FMM95]). Le principal outil pour prouver cette analyticité est la fonction de Lyapunov dans le critère d'ergodicité de Foster. Nous obtenons une fonction de Lyapunov quadratique adaptée. Mais la dynamique de notre modèle ne permet pas d'appliquer les résultats de [MM79, FMM95] en

raison de la politique du choix local qui implique un phénomène de déposition. Cette question n'est pas tranchée et reste pour le moment inachevée. Sous cette hypothèse, on obtient le résultat suivant.

Proposition 1.6.1

Pour le modèle de choix local avec $N \geqslant 3$, la probabilité stationnaire $\pi_m(\rho)$ qu'une file d'attente donnée ait m clients est telle que

$$\pi_m(\rho) = 12(\rho/2)^{2m-1} + O(\rho^{2m}) \text{ pour } \rho \text{ au voisinage de } 0.$$

Comme prévu, les performances de la politique du choix local se situent entre les deux autres politiques. Cependant, pour un trafic faible, son comportement est plus proche de l'absence de choix que du choix aléatoire. En effet, les deux premières asymptotiques sont exponentielles tandis que la troisième est doublement exponentielle en ρ.
Les asymptotiques du trafic léger obtenues dans [DFM18] concernent la limite lorsque t tend vers l'infini d'abord et ensuite N alors que dans l'approche champ-moyen pour le modèle du choix aléatoire, la limite est établie lorsque N d'abord et ensuite t tendent vers l'infini. La comparaison des performances des deux modèles est justifiée par la possibilité de permuter l'ordre les deux limites, voir [VDK96] pour plus de détails.

Chapitre 2

Du vélopartage à l'autopartage : grand système avec forte interaction

Collaborations avec Cédric Bourdais, Christine Fricker, Bianca Marin Moreno, Teodora Pescu, Amaury Philippe, Martin Trépanier et Florian Verdier.

2.1 Motivation

Du vélopartage vers l'autopartage, quelle différence majeure ? C'est la réservation : de la voiture, de la place de parking, ou des deux. Cet aspect inexistant en vélopartage est simple à modéliser, ne casse pas le caractère Markovien du processus d'état du système et reste analysable par l'approche champ-moyen avec une interprétation files d'attente du point d'équilibre du système dynamique limite. Cependant, le cas le plus intéressant à mon sens est celui de la double réservation où l'usager réserve à la fois la voiture et la place de parking, typiquement ce que proposait Autolib', le grand système d'autopartage parisien avant son arrêt en 2017. La modélisation d'un tel système réel en tant que grand système avec forte interaction nous fournit un exemple similaire au célèbre modèle de Gibbens, Hunt et Kelly [GHK90] présentant des interactions fortes qui disparaissent dans la limite champ-moyen. La technique utilisée est spécifique à la double réservation comparée au cas de la simple réservation.

2.2 Simple réservation

Il s'agit de la réservation de la voiture seule, ou de la place de parking seule. Pour plus de clarté, je considère ici la réservation de la voiture sans possibilité de réserver la place de parking. Je présenterai trois modèles, le premier que je qualifierai de modèle de base, le second plus réaliste et le troisième avec boucles. D'un point de vue applicatif, réserver la voiture sans possibilité de réserver la place de parking a du sens dans des systèmes d'autopartage sans stations physiques, appelés *free-floating*. En effet, la dernière décennie a vu l'apparition de systèmes de partages de voitures où l'usager peut récupérer et garer ces voitures partagées n'importe où dans l'espace public à l'intérieur d'une zone géographique délimitée, appelée zone de service. Une première approche pour analyser ces systèmes a été de les assimiler à un système d'autopartage classique en découpant la zone de service en des petites zones de l'ordre de $1\,km^2$ considérées comme des stations dont la capacité fixe était choisie arbitrairement (par exemple le nombre maximum de voitures partagées garées dans la zone considérée). Voir la Figure 2.1. La modélisation proposée dans ce chapitre suivra cette approche et on considère alors un système d'autopartage classique avec des stations physiques où seule la réservation de la

voiture est possible. Une seconde approche spécifique au *free-floating* sera présentée au chapitre suivant.

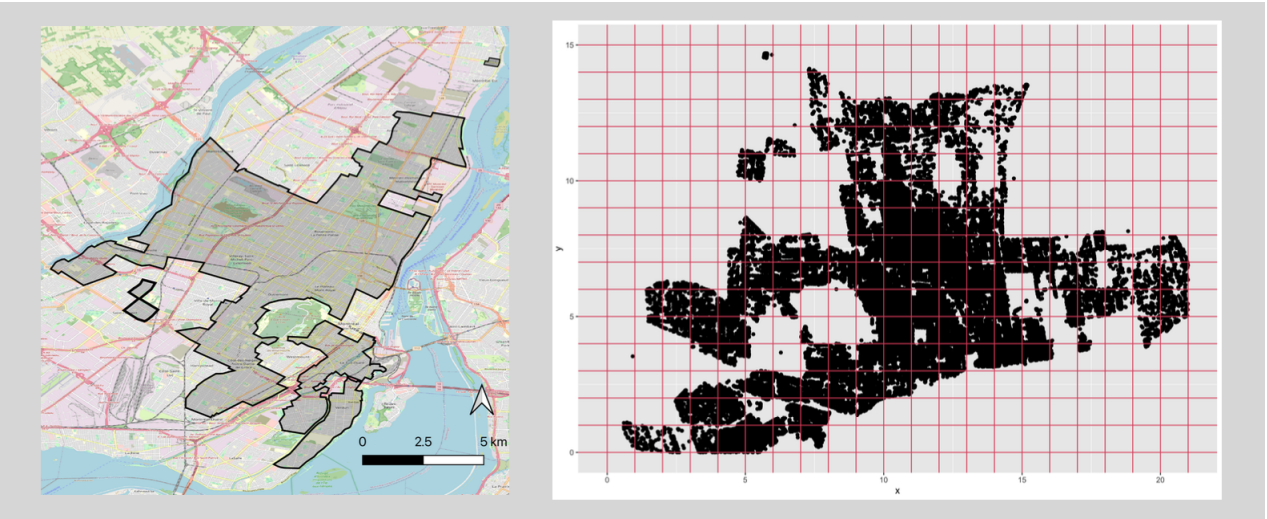

Figure 2.1 – A gauche : zone de service *free-floating* $\sim 150\,km^2$de Communauto à Montréal. A droite : reconstitution de la zone de service à l'aide des coordonnées GPS des voitures partagées. Découpage de la zone de service en une grille de petites zones de $1\,km^2$.

Le modèle de base

Ce premier modèle est un modèle homogène où toutes les stations ont les mêmes propriétés. En particulier, le comportement des usagers est le même quelle que soit la station d'où ils partent et vers laquelle ils se dirigent. Les dynamiques sont les suivantes.

- L'arrivée d'un usager à une station donnée avec une voiture disponible marque le début de sa réservation. Si aucune voiture n'est disponible, l'usager quitte le système.
- À la fin de sa réservation, l'usager prend la voiture pour commencer un trajet vers une station destination, choisie au hasard parmi toutes les stations de la zone de service, y compris la station d'où il vient.
- A la fin de son trajet, l'usager gare la voiture dans la station de destination préalablement choisie, s'il y a une place de parking disponible dans cette station. Sinon, il choisit une autre station au hasard, effectue un trajet et gare la voiture si une place de parking est disponible. Il répète cette opération jusqu'à ce qu'il puisse garer la voiture.

Les variables aléatoires suivantes sont indépendantes avec une distribution exponentielle : temps d'inter-arrivée, temps de réservation et temps de trajet. Ceci permet de traiter des processus de Markov discrets, qui sont faciles à manipuler. Ce modèle est une vue simplifiée du système réel d'autopartage en *free-floating* où l'usager doit réserver sa voiture avant un trajet sans possibilité d'annulation après la réservation. Mais il s'agit de la première étape pour construire et comprendre des modèles plus complexes présentés plus loin dans ce chapitre. En plus des notations introduites pour le modèle homogène sans réservation (Vélib'), on introduit :

- η le paramètre de la distribution exponentielle des durées de réservation. La durée moyenne de réservation est $1/\eta$.
- $V_i^N(t)$ le nombre de voitures disponibles et $R_i^N(t)$ celui des voitures réservées dans la station i à l'instant $t \geqslant 0$. L'état des N stations à l'instant $t \geqslant 0$ est noté par $(V_i^N(t), R_i^N(t), 1 \leq i \leq N)$.

Le processus mesure empirique associé

$$Y_{k,l}^N(t) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{(V_i^N(t)=k, R_i^N(t)=l)}$$

est un processus de Markov irréductible sur un espace d'états fini. Il converge, quand N devient grand, vers un système dynamique de la forme

$$y'(t) = y(t)\, L_{y(t)}$$

où $L_{y(t)}$ est le générateur infinitésimal du nombre de clients, noté $(V(t), R(t))$ dans un tandem de deux files contraintes par une capacité totale c, de taux d'arrivée $\mu(s - \sum_{k+l\leq c} y_{k,l})(t)$ et tel que

— la première file est une file $M/M/1$ de taux de service λ,

— la deuxième file est une file $M/M/\infty$ de taux de service η.

Une illustration de la file typique est donnée par la figure 2.2.

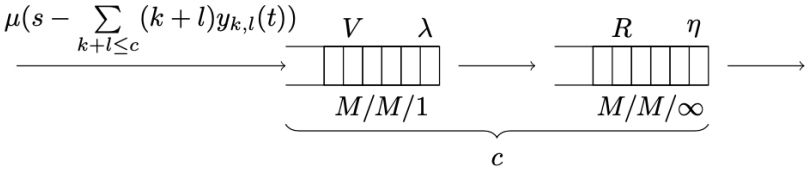

Figure 2.2 – Une station typique est un tandem de deux files d'attente avec capacité globale c.

Notons que le taux d'arrivée peut être réécrit $\mu\big(s - \mathbb{E}(V(t) + R(t))\big)$. On peut prouver que le processus de Markov inhomogène limite $(V(t), R(t))$ existe. Il est appelé processus de McKean–Vlasov. La preuve est technique mais standard. Elle sera reprise dans un cadre plus général, celui de la double réservation. Comme dans le premier chapitre, un point clé de cette interprétation files d'attente est l'expression explicite de la mesure invariante associée à ce tandem dont l'existence et l'unicité n'est donc pas évidente. Ceci est donné par le résultat suivant.

Théorème 2.2.1 (*Existence et unicité de la mesure invariante*)

Le processus limite $(V(t), R(t))$ admet une unique mesure invariante π sur $\{(k,l) \in \mathbb{N}^2,\ k+l \leq c\}$ définie par

$$\pi(k,l) = \frac{1}{Z}\rho_V^k \frac{\rho_R^l}{l!} \tag{2.1}$$

où Z est la constante de normalisation, les variables ρ_V et ρ_V sont solutions du système

$$(S_1): \begin{cases} \rho_R = \dfrac{\lambda}{\eta}\rho_V \\ \rho_V = \dfrac{\mu}{\lambda}\Big(s - \sum_{k+l\leq c}(k+l)\pi(k,l)\Big). \end{cases}$$

Notons d'abord que la première équation du système (S_1) permet de réduire le nombre d'inconnues. On peut alors exprimer la constante de normalisation Z en fonction d'une seule variable, ρ_V par exemple. Ainsi, la mesure invariante π est une fonction explicite de ρ_V. La seconde équation du système (S_1) devient donc une équation de point fixe en ρ_V. La preuve de l'existence est classique, où un bon candidat pour être cette mesure invariante est la troncature de la mesure invariante d'un tandem similaire de deux files non contraintes par une capacité finie. En effet, il est bien connu depuis

[BCMP75] qu'un tel tandem à capacité infinie admet une unique mesure invariante de la forme (2.1). L'unicité est une conséquence de la monotonie puisque, à s fixé, les deux termes du membre de droite de la relation

$$s = \frac{\lambda}{\mu}\rho_V + \frac{1}{Z}\sum_{k+l\leq c}(k+l)\rho_V^k\frac{\rho_R^l}{l!}$$

sont strictement croissants en ρ_V.

Un modèle plus réaliste

Afin de modéliser un système d'autopartage plus réaliste, considérons un modèle où la réservation de la voiture est possible mais non obligatoire, l'annulation de trajet et les réservations enchaînées sont autorisées. Introduisons les paramètres suivants.

- α est la probabilité de réserver une voiture au départ,
- β est la probabilité d'annuler le trajet à la fin de la réservation,
- β' est la probabilité d'enchaîner une autre réservation suite à la première réservation.

Les dynamiques sont alors modifiées comme suit.

- Les usagers arrivent à une station selon un processus de Poisson. Si une voiture est disponible, soit il la réserve avec une probabilité α, soit il la prend sans la réserver avec une probabilité $1-\alpha$. S'il n'y a pas de voiture disponible, il quitte le système.
- À la fin de réservation, avec une probabilité β, l'usager ne prend pas la voiture et quitte le système. Dans le cas contraire, deux possibilités s'offrent à l'usager : soit il effectue une nouvelle réservation de la même voiture avec la probabilité $\beta' < 1-\beta$, soit avec probabilité $1-\beta-\beta'$, il prend la voiture pour commencer son trajet vers sa destination, choisie au hasard.
- La fin du trajet est la même que pour le modèle de base précédemment décrit.

Le résultat suivant donne le comportement limite à grande échelle d'une station donnée.

Théorème 2.2.2 (*Convergence champ-moyen*)

Quand N tend vers l'infini, avec M/N qui tend vers une constante s, le processus d'état $(V_i^N(t), R_i^N(t))$ d'une station donnée i converge vers $(V(t), R(t))$ dont la distribution $(y(t))$ est celle du nombre de clients dans deux files contraintes par une capacité totale c et formant un réseau de Jackson ouvert telles que :

- *la première file est une une file $M/M/1$ avec taux de service λ,*
- *la deuxième file est une file $M/M/\infty$ avec taux de service η,*
- *le taux d'arrivée à la première file est $\mu(s - \sum_{k+l\leq c} y_{k,l}(t))$,*
- *pas d'arrivée à la deuxième file.*

Après être servi à une des deux files $i \in \{1,2\}$, le client est envoyé à la file $j \in \{1,2\}$ avec probabilité $p_{i,j}$ telle que

$$p_{1,1} = 0, \quad p_{1,2} = \alpha, \quad p_{2,1} = \beta, \quad p_{2,2} = \beta'$$

ou bien quitte le système. Une illustration de cette file typique est donnée par la Figure 2.3

L'existence et l'unicité de la mesure invariante est similaire au modèle de base et on obtient une mesure invariante π de la même forme que (2.1), caractérisée par deux paramètres ρ_R et ρ_V solutions du système suivant :

$$(S_2): \begin{cases} \rho_R = \dfrac{\lambda\alpha}{\eta(1-\beta')}\rho_V \\ \rho_V = \dfrac{\mu(1-\beta')}{\lambda(1-\beta'-\alpha\beta)}\Big(s - \sum_{k+l\leq c}(k+l)\pi(k,l)\Big). \end{cases}$$

Notons la similarité entre les deux systèmes (S_1) et (S_2) caractérisant le point d'équilibre du système dynamique pour les deux modèles. Cette similarité peut être interprétée comme une équivalence entre le modèle de base et le modèle raffiné. En effet, il n'est pas difficile de voir que le second modèle (avec paramètres λ, η, μ, α, β et β') est équivalent au modèle de base (avec paramètres λ, $\tilde{\eta}$ et $\tilde{\mu}$), en terme de mesure invariante du système limite, et ce en posant :

$$\begin{cases} \tilde{\mu} = \dfrac{\mu(1-\beta')}{\lambda(1-\beta'-\alpha\beta)} \\ \tilde{\eta} = \dfrac{\eta(1-\beta')}{\alpha}. \end{cases}$$

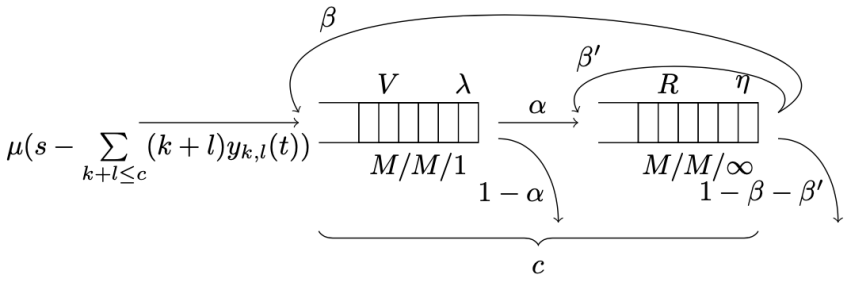

Figure 2.3 – Une station typique est réseau de Jackson ouvert composé de deux files d'attente avec une capacité globale c.

Un modèle avec boucles

En analysant les données, il s'avère qu'une proportion importante de déplacements est en boucle (origine = destination). Cela ne correspond pas aux modèles précédents où les trajets en boucles représentent une proportion $1/N$ des déplacements, ce qui est négligeable lorsque N devient grand. Le modèle suivant est introduit pour prendre en compte cet aspect. Pour éviter les notations fastidieuses et se concentrer sur la question des boucles, le modèle proposé est une variante du modèle de base avec quelques notations supplémentaires.

— p est la probabilité limite, lorsque N devient grand, qu'un trajet soit une boucle.

— μ_L est le paramètre de la distribution exponentielle de la durée d'un trajet en boucle.

En effet, les dynamiques du modèle de base sont modifiées comme suit. À la fin de la réservation, l'usager prend la voiture réservée pour un trajet en boucle avec la probabilité p. Ce trajet dure un temps aléatoire dont la distribution est exponentielle de paramètre μ_L, qui peut être distinct de μ. Dans le cas contraire, l'usager choisit une destination au hasard parmi toutes les stations, y compris la station d'où il vient. Il en résulte que la probabilité qu'un trajet soit une boucle est $p + (1-p)/N$. Le processus décrivant l'état d'une station donnée i est alors à trois composantes $(V_i^N(t), R_i^N(t), L_i^N(t))$ où $L_i^N(t)$ est le nombre de voitures effectuant un trajet en boucle de la station i vers elle même. La limite champ-moyen est similaire, avec une écriture en terme d'un réseau de Jackson ouvert composé de trois files d'attente.

Théorème 2.2.3 (*Convergence champ-moyen*)

Quand N tend vers l'infini, avec M/N tend vers une constante s, le processus d'état $(V_i^N(t), R_i^N(t), L_i^N(t))$ d'une station donnée i converge vers $(V(t), R(t), L(t))$ dont la distribution $(y(t))$ est celle d'un réseau de Jackson ouvert composé de trois files d'attente telles que :

— *les deux premières files sont contraintes par une capacité totale* c,
— *la première file est une une file* $M/M/1$ *avec taux de service* λ,
— *la deuxième file est une file* $M/M/\infty$ *avec taux de service* η,
— *la troisième file est une file* $M/M/\infty$ *avec taux de service* μ_L,
— *le taux d'arrivée à un instant* $t \geqslant 0$ *à la première file est*

$$\mu(s - \sum_{m\geqslant 0,\, k+l\leq c} (k+l+m) y_{k,l,m}(t)),$$

— *pas d'arrivée aux deux autres files.*
— *pour* $i, j \in \{1,2,3\}$, *les probabilités* $p_{i,j}$ *sont données par*

$$p_{1,2} = 1, \quad p_{2,3} = p, \quad p_{3,1} = 1, \quad p_{i,j} = 0 \text{ sinon.}$$

Une illustration de cette file typique est donnée par la Figure 2.4

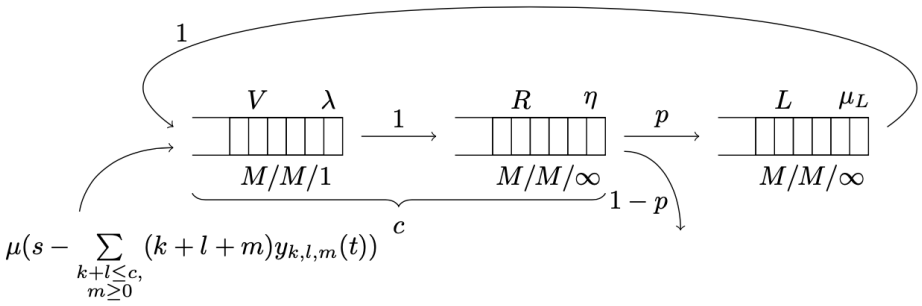

Figure 2.4 – Une zone typique est un tandem de deux files d'attente avec une capacité globale de c.

Il est à noter que le taux d'arrivée peut être réécrit comme

$$\mu(s - \sum_{m\geqslant 0,\, k+l\leq c} (k+l+m) y_{k,l,m}(t)) = \mu(s - \mathbb{E}(V(t) + R(t) + L(t))).$$

Bien que le système caractérisant le point d'équilibre de ce système dynamique soit de taille 3, il reste simple à manipuler

$$(S_3) : \begin{cases} \rho_R = \dfrac{\lambda}{\eta}\rho_V \\ \rho_L = \dfrac{p\lambda}{\mu_l}\rho_V \\ \rho_V = \dfrac{\mu}{\lambda(1-p)}\Big(s - \sum_{m\geqslant 0,\, k+l\leq c}(k+l+m)\pi(k,l,m)\Big). \end{cases}$$

L'existence et l'unicité de la mesure invariante est similaire aux deux modèles précédents et on obtient une mesure invariante π de forme produit sur $\{(k,l,m) \in \mathbb{N}^3,\, k+l \leqslant c\}$ telle que

$$\pi(k,l,m) = \frac{1}{Z}\rho_V^k \frac{\rho_R^l}{l!}\frac{\rho_L^m}{m!}$$

où les trois paramètres ρ_R, ρ_V et ρ_L sont solutions du système (S_3). L'état limite d'une station du système est donné par le couple $(V(t), R(t))$ dont la mesure invariante $\pi_{k,l}$ sur $\{(k,l) \in \mathbb{N}^2,\, k+l \leq c\}$

est donnée, avec abus de notations, par la même forme produit (2.1), avec

$$\begin{cases} \rho_R = \dfrac{\lambda}{\eta}\rho_V \\ \rho_V = \dfrac{1}{\lambda(\dfrac{p}{\mu_L} + \dfrac{1-p}{\mu})}\Big(s - \sum_{k+l\leq c}(k+l)\pi(k,l)\Big). \end{cases}$$

Il est simple de voir que si les durées des trajets en boucle ne diffèrent pas en terme de loi des autres trajets, ce qui implique que $\mu_L = \mu$, alors on retrouve le système (S_1) et donc le même comportement limite donné par le modèle de base. Dans le cas contraire si $\mu_L \neq \mu$, alors il suffit de remplacer le temps moyen d'un trajet $1/\mu$ dans le modèle de base par $(\dfrac{p}{\mu_L} + \dfrac{1-p}{\mu})$ pour retrouver un comportement limite équivalent entre les deux modèles. Ceci montre bien la robustesse du modèle de base qui intègre différentes variantes suivant le besoin de tenir compte d'aspects pratiques de l'application considérée ici à savoir l'autopartage.

Etude de performance

La force des modèles homogènes est de donner des résultats explicites sur le comportement limite du système. En effet, l'expression de la limite champ-moyen à l'équilibre est suffisamment simple pour être utilisée. L'objectif ultime est de présenter des résultats qualitatifs et quantitatifs sur le comportement du système à grande échelle en temps long et d'étudier l'influence de la réservation. Afin de faire une comparaison entre les deux modèles, avec et sans réservation, fixons d'abord le choix de la mesure de performance. Rappelons que le problème principal est la présence de stations dites *problématiques*, celles qui n'ont pas de voitures disponibles ou de places de parking disponibles. Il est alors naturel de considérer comme métrique la proportion limite Pb de stations problématiques à l'équilibre. Elle est donnée par

$$Pb = \pi_{0,\cdot} + \pi_S - \pi_{0,c} \tag{2.2}$$

où π est la distribution conjointe limite stationnaire des voitures disponibles et réservées dans une station donnée,

$$\begin{cases} \pi_{0,\cdot} = \pi_{0,0} + \ldots + \pi_{0,c}, & \text{proportion limite des stations sans voiture disponible} \\ \pi_S = \pi_{0,c} + \pi_{1,c-1} + \ldots + \pi_{c,0}, & \text{proportion limite des stations saturées, sans place disponible.} \end{cases}$$

Pour éviter des notations trop lourdes, on se limitera au modèle de base où ρ_V sera désigné simplement par ρ. En effet, l'équation du point fixe, vérifiée par ρ pour le modèle de base, est valable, avec de simples adaptations, pour le modèle raffiné et le modèle avec boucles. Rappelons que, d'après le Théorème 2.2.1, la mesure stationnaire π est le produit d'une distribution géométrique de paramètre ρ et d'une distribution de Poisson de paramètre $\rho\lambda/\eta$ sur $\{(k,l) \in \mathbb{N}^2, k+l \leqslant c\}$, où ρ est donné par une équation de point fixe. Cette métrique est utilisée dans les travaux connexes pour de tels systèmes [FG16], [FGM12] et [FMB20]. Elle facilitera donc la comparaison. Comme dans ces travaux, notre objectif est d'étudier l'influence du dimensionnement sur les performances du système, l'ajustement de la taille de la flotte étant le premier levier de l'opérateur. L'intuition est qu'il existe une taille de flotte optimale caractérisée par un ratio $s = \lim_{N\to+\infty} M/N$ tel que la probabilité qu'une station soit problématique soit minimale. Par la courbe paramétrique $\rho \mapsto (s(\rho), Pb(\rho))$, il en résulte que la proportion limite de stations problématiques Pb est aussi fonction de s puisque s et Pb sont des fonctions explicite de ρ. L'étude de cette courbe paramétrique établit l'existence et l'unicité d'une taille de flotte optimale s^*. L'écart par rapport à l'optimum pour un modèle homogène sans

réservation, typiquement un système de vélopartage tel que étudié dans [FG16] peut être mesuré. En effet, il est prouvé dans [FG16] que la proportion limite stationnaire de stations problématiques a un minimum unique $Pb^* = 2/(c+1)$ pour le paramètre de taille de la flotte $s^* = c/2 + \lambda/\mu$. Pour le modèle de base avec réservation, on obtient que

$$\begin{cases} \lim\limits_{\eta\to+\infty} \dfrac{Pb^* - 2/(c+1)}{\lambda/\eta} = \dfrac{2}{(c+1)^2} \\ \lim\limits_{\eta\to+\infty} \dfrac{s^* - (c/2+\lambda/\mu)}{\lambda/\eta} = \dfrac{c^2-3c/2-1}{2(c-1)^2}. \end{cases}$$

Notons que, si c tend vers l'infini, ces limites sont respectivement équivalentes à $2/c^2$ et $1/2$. En conclusion, Pb et s à l'optimalité ont respectivement des termes supplémentaires en λ/η par rapport au cas sans réservation de voiture et sont donnés pour c grand par

$$Pb^* \sim \frac{2}{c+1} + \frac{2}{c^2}\frac{\lambda}{\eta} \text{ pour } s^* \sim \frac{c}{2} + \frac{\lambda}{\mu} + \frac{\lambda}{2\eta}.$$

Ces résultats quantitatifs confirment l'impact "négatif" de la réservation sur le système puisque la proportion limite minimale de stations problématiques est plus grande que dans le cas sans réservation. Ceci doit être tempéré par le fait que le terme additif $2\lambda/c^2\eta$ est petit pour un temps de réservation moyen $1/\eta$ assez petit. On voit aussi que l'opérateur aura intérêt à "mettre" plus de voitures par station pour tenir compte de la réservation de la voiture.

La courbe paramétrique $\rho \mapsto (s(\rho), Pb(\rho))$ est représentée dans la Figure 2.5 pour des paramètres inférés à partir des données réelles fournies par l'opérateur Communauto pour son système d'autopartage en *free-floating* à Montréal, appelé aussi *Flex* : les temps de trajet ont une distribution exponentielle avec une moyenne de 1 h 22 mn et les temps de réservation ont une distribution exponentielle avec une moyenne de 17 mn. Ces moyennes sont issues des données bien que les distributions des temps de trajet et de réservation réelles ne soient pas exponentielles. L'usager réserve avec une probabilité $\alpha = 0,84$ et quitte le système à la fin de réservation sans prendre la voiture avec une probabilité $\beta = 0,2$. La Figure 2.5 confirme que l'écart en terme de performance dû à la réservation est assez négligeable. On constate que la proportion limite Pb est élevée dans deux cas :

— dans n'importe quel régime (faible ou fort trafic), lorsque le nombre de voitures est faible, et Pb augmente avec un trafic élevé,

— en cas de faible trafic, lorsque le nombre de voitures par stations est proche de la capacité.

Le premier scénario peut être évité en augmentant le nombre de voitures partagées et le second en le réduisant par rapport à la capacité.

Nous observons sur le graphique l'existence et l'unicité du minimum intuitivement deviné. Notons que le minimum se déplace vers la droite avec la demande, ce qui signifie que le paramètre de taille de flotte optimale noté s^* est croissant avec la demande dont le taux est noté λ. Ainsi, avec une forte demande, un système dont le paramètre de taille de flotte s est proche de c peut être performant.

La réservation de la place de parking

D'un point de vue applicatif, un modèle où seule la place de parking peut être réservée par l'usager ne correspond pas à un système d'autopartage réel. Mais, d'un point de vue technique, c'est un modèle intermédiaire qui permet d'aborder les questions d'existence et d'unicité du point d'équilibre du système dynamique associé au modèle de la double réservation présenté à la prochaine section. Pour ce modèle de réservation de la place de parking, on obtient dans [FMB20] que l'état d'une station donnée se comporte en approximation champ-moyen comme le nombre de clients dans deux files d'attente en tandem dont la capacité globale est c : une file $M/M/\infty$ (les réservations), avec

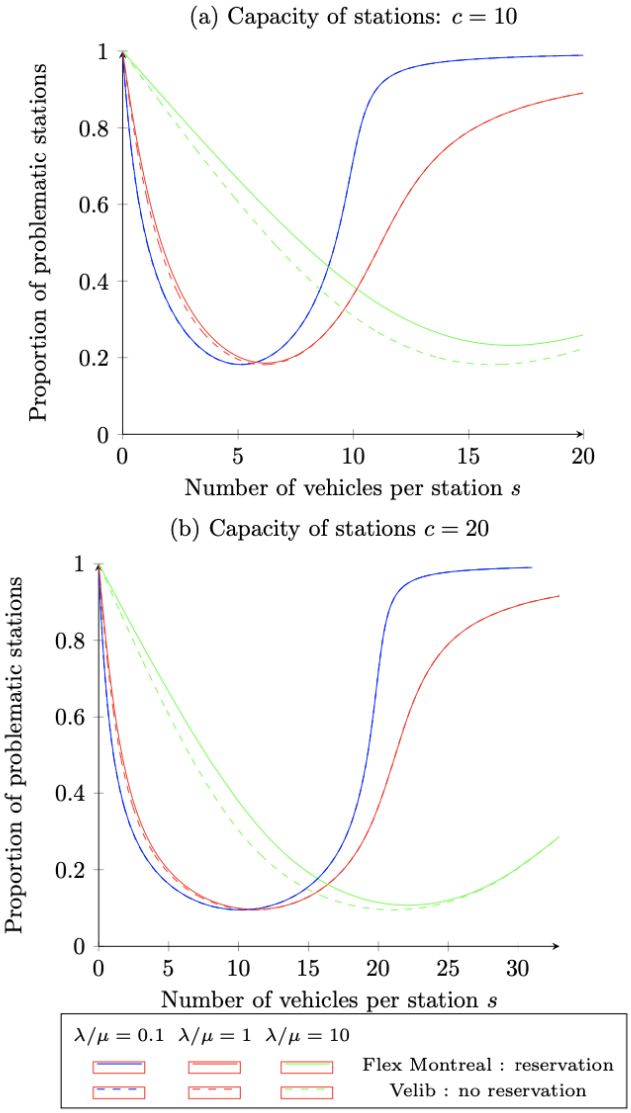

Figure 2.5 – Courbe paramétrique $\rho \mapsto (s(\rho), Pb(\rho))$ pour $c = 10$ **(a)** et $c = 20$ **(b)** pour le modèle d'autopartage avec réservation de la voiture (Flex Montréal) et celui sans réservation (vélopartage). Pour les deux cas : $\nu/\mu = 4$.

un taux d'arrivée $\lambda(1 - y_{\cdot,0})$, où $y_{\cdot,0}$ est le proportion limite de stations sans voitures, et un taux de service μ, et une file $M/M/1$ (les voitures), où les clients viennent de la première file, avec un taux de service $\lambda(1 - y_S)$ où y_S est la proportion de stations saturées, c'est-à-dire sans place de parking disponible. Un point surprenant est que, contrairement à la réservation de la voiture, établir l'unicité du point d'équilibre ici est plus compliqué d'un point de vue combinatoire vue la relation implicite entre les deux paramètres ρ_R et ρ_V.

2.3 Double réservation

La réservation devient techniquement un vrai défi quand on tient compte de la réservation de la voiture (au départ) et de la place de parking (à destination). C'est typiquement ce que proposait Autolib',

le système d'autopartage qui a existé à Paris de 2011 à 2018. Sa flotte était composée, en juillet 2016, de 3 980 véhicules électriques, appelés Bluecars, répartis dans 1 084 stations de l'agglomération parisienne avec 5 935 points de charge. Plus de 126 900 abonnés avaient été enregistrés pour le service. Intuitivement, la double réservation est un confort pour les usagers, qui a forcément un coût pour le système en terme de performance. Dans [FM25], nous proposons un modèle homogène afin de mesurer l'impact de la double réservation sur le comportement en temps long du système.

Le modèle

Dans [FM25], on considère un modèle où l'usager réserve une voiture à une station de départ donnée et, en même temps, une place de parking à une station destination. S'il n'y a pas de voiture ou de place de parking disponible, l'usager quitte le système. Dans le cas contraire, il s'écoule un certain temps, appelé temps de réservation, entre la réservation de la voiture et sa prise en charge. Le temps de réservation a une distribution exponentielle avec une moyenne de $1/\nu$. Vient ensuite le trajet dont la durée est supposée avoir une distribution exponentielle avec une moyenne de $1/\mu$. A la fin de son trajet, l'usager gare la voiture à l'emplacement réservé dans la station de destination et quitte le système.
Un espace d'état tridimensionnel (voitures réservées, places réservées et voitures disponibles) pour chaque station ne permet pas de décrire complètement le système comme un processus de Markov, car le temps de réservation de la place de parking est la somme de deux variables à distribution exponentielle (temps de prise en charge de la voiture et temps de trajet). Il faut alors distinguer les places de parking réservées par les usagers qui ne se déplacent pas encore de celles réservées par les usagers qui se déplacent. Pour une description markovienne du système, il faut considérer un processus d'état qui tient compte de l'interaction entre les stations départ et destination $\left(X^N(t) = X^N_{i,j}(t), 0 \leq i,j \leq N\right)$ où, pour $1 \leq i,j \leq N$, à l'instant t,

- $X^N_{i,j}(t)$ est le nombre de voitures réservées à la station i avec un espace de stationnement réservé à la station$j$,
- $X^N_{0,j}(t)$ est le nombre de places de parking réservées à la station j par des usagers conduisant,
- $X^N_{i,0}(t)$ est le nombre de voitures disponibles à la station i.

Le processus $(X^N(t))$ est un processus de Markov irréductible sur un espace d'états fini

$$\left\{ x = (x_{i,j}) \in \mathbb{N}^{(N+1)^2}, \sum_{j=0}^{N} x_{i,j} + \sum_{j=0}^{N} x_{j,i} \leqslant c, \sum_{0 \leqslant i,j \leqslant N}^{N} x_{i,j} = M \right\}, \tag{2.3}$$

où c désigne la capacité finie de chaque station.

A la recherche d'un "bon" processus d'état

La dimension de $(X^N(t))$ est de l'ordre de N^2, ce qui ne se prête pas à une analyse champ-moyen lorsque N tend vers l'infini. Nous introduisons alors un autre processus, $(Z^N(t) = Z^N_i(t),\ 1 \leq i \leq N)$ tel que l'état d'une station i à un instant $t \geqslant 0$ est donné par $\left(Z^N_i(t) := R^{r,N}_i(t), R^N_i(t), V^N_i(t), V^{r,N}_i(t)\right)$ où

- $R^{r,N}_i(t)$ est le nombre de places de parking réservées par des usagers qui n'ont pas commencé leurs trajets,
- $R^N_i(t)$ est le nombre de places de parking réservées par des usagers en déplacement,
- $V^N_i(t)$ est le nombre de voitures disponibles,
- $V^{r,N}_i(t)$ est le nombre de voiture réservées.

Le processus $(Z^N(t))$ permet de bien décrire le système et de capturer sa performance, comme le fait qu'une station soit vide ou pleine. Ainsi, il donne un état suffisamment fin des stations tout en gardant une dimension linéaire en N. Il s'exprime comme fonction du processus de Markov $(X^N(t))$

$$R_i^N(t) = X_{0,i}^N(t), \qquad R_i^{r,N}(t) = \sum_{j=1}^N X_{j,i}^N(t),$$
$$V_i^N(t) = X_{i,0}^N(t), \qquad V_i^{r,N}(t) = \sum_{j=1}^N X_{i,j}^N(t).$$

Néanmoins, le processus $(Z^N(t))$ n'est pas markovien. Pour s'en convaincre, il suffit d'écrire ses équations d'évolutions et de voir qu'elles ne sont pas autonomes, c'est-à-dire qu'elles ne s'écrivent pas exclusivement en fonction du processus $(Z^N(t))$. Pour celà, introduisons les processus ponctuels indépendants suivants. La réservation d'une voiture à la station i est un point d'un processus de Poisson $\mathcal{N}_{\lambda,i}^{U,N}$ sur $\mathbb{R}_+^2$. Pour la réservation d'une voiture à la station i et d'une place de parking à la station j, le temps écoulé entre le moment où l'utilisateur effectue une réservation et le moment où la voiture est récupérée est associé à un processus de Poisson $\mathcal{N}_{\nu,i,j}$ sur $\mathbb{R}_+$. Nous avons besoin d'une suite i.i.d. de tels processus $(\mathcal{N}_{\nu,i,j,l},\, l \in \mathbb{N})$ car les voitures sont prises indépendamment les unes des autres. Et de même pour les durées de trajet des voitures déposées à la station j, associées à une suite $(\mathcal{N}_{\mu,j,l},\, l \in \mathbb{N})$ de processus de Poisson indépendants. Ainsi, le processus $(Z^N(t))$ est solution des équations d'évolution stochastiques suivantes.

$$\begin{cases} dR_i^N(t) = & \sum_{l=1}^\infty \sum_{j=1}^N \mathbf{1}_{\{l \leqslant X_{j,i}^N(t^-)\}} \mathcal{N}_{\nu,j,i,l}(dt) - \sum_{l=1}^\infty \mathbf{1}_{\{l \leqslant R_i^N(t^-)\}} \mathcal{N}_{\mu,i,l}(dt), \\ dR_i^{r,N}(t) = & \sum_{j=1}^N \mathbf{1}_{\{V_j^N(t^-)>0,\, S_i^N(t^-)<c\}}\, \mathcal{N}_{\lambda,j}^{U,N}(dt,\{i\}) - \sum_{j=1}^N \left(\sum_{l=1}^\infty \mathbf{1}_{\{l \leqslant X_{j,i}^N(t^-)\}} \mathcal{N}_{\nu,j,i,l}(dt)\right), \\ dV_i^N(t) = & \sum_{l=1}^\infty \mathbf{1}_{\{l \leqslant R_i^N(t^-)\}} \mathcal{N}_{\mu,i,l}(dt) - \sum_{j=1}^N \mathbf{1}_{\{V_i^N(t^-)>0,\, S_j^N(t^-)<c\}}\, \mathcal{N}_{\lambda,i}^{U,N}(dt,\{j\}), \\ dV_i^{r,N}(t) = & \sum_{j=1}^N \mathbf{1}_{\{V_i^N(t^-)>0,\, S_j^N(t^-)<c\}}\, \mathcal{N}_{\lambda,i}^{U,N}(dt,\{j\}) - \sum_{j=1}^N \left(\sum_{l=1}^\infty \mathbf{1}_{\{l \leqslant X_{i,j}^N(t^-)\}} \mathcal{N}_{\nu,i,j,l}(dt)\right) \end{cases} \tag{2.4}$$

où on désigne par

$$S_i^N(t) = R_i^{r,N}(t) + R_i^N(t) + V_i^N(t) + V_i^{r,N}(t)$$

le nombre de places de parking indisponibles à la station i à l'instant t. Notons que $S_i^N(t)$ est une fonction de $Z_i^N(t)$.
Il est remarquable que, malgré cette description non-markovienne, l'état $(Z_i^N(t))$ d'une station i donnée $(1 \leq i \leq N)$ converge en distribution, lorsque N tend vers l'infini, vers un processus de Markov inhomogène $(\overline{Z}(t)) = (\overline{R^r}(t), \overline{R}(t), \overline{V}(t), \overline{V^r}(t))$ satisfaisant l'équation de Fokker-Planck suivante

$$\begin{aligned} \frac{d}{dt}\mathbb{E}\left(f(\overline{Z}(t))\right) &= \lambda \mathbb{P}(\overline{V}(t)>0)\mathbb{E}\left((f(\overline{Z}(t)+e_1) - f(\overline{Z}(t))1_{\{\bar{S}(t)<K\}}\right) \\ &+ \nu \mathbb{E}\left((f(\overline{Z}(t)+e_2-e_1) - f(\overline{Z}(t)))1_{\{\overline{R^r}(t)>0\}}\right) \\ &+ \mu \mathbb{E}\left((f(\overline{Z}(t)+e_3-e_2) - f(\overline{Z}(t)))1_{\{\overline{R}(t)>0\}}\right) \\ &+ \lambda \mathbb{P}(\overline{S}(t)<K)\mathbb{E}\left((f(\overline{Z}(t)+e_4-e_3) - f(\overline{Z}(t)))1_{\{\overline{V}(t)>0\}}\right) \\ &+ \nu \mathbb{E}\left((f(\overline{Z}(t)-e_4) - f(\overline{Z}(t)))1_{\{\overline{V^r}(t)>0\}}\right) \end{aligned} \tag{2.5}$$

avec $e_1 = (1,0,0,0)$, $e_2 = (0,1,0,0)$, $e_3 = (0,0,1,0)$, $e_4 = (0,0,0,1)$, f une fonction à support fini dans $\mathbb{N}^4$ et $\overline{S}(t) = \overline{R^r}(t) + \overline{R}(t) + \overline{V}(t)) + \overline{V^r}(t)$ le nombre limite des places de parking indisponibles à la station i à l'instant t. Le processus asymptotique $(\overline{Z}(t)) = ((\overline{R^r}(t), \overline{R}(t), \overline{V}(t), \overline{V^r}(t))$ est un processus de sauts avec des taux dépendants du temps, appelé processus de McKean–Vlasov. Cette convergence champ-moyen n'est pas du tout standard puisque la limite champ-moyen est généralement obtenue pour un processus de Markov. Le même phénomène est prouvé dans un cadre entièrement différent, pour un modèle de réseau avec des défaillances dans [ARS18]. Il est à noter que, contrairement au modèle de la double réservation étudié dans [FM25], le modèle considéré dans [ARS18] induit des sauts simultanés, ce qui rend les preuves encore plus techniques. Comme le célèbre modèle de Gibbens, Hunt et Kelly ([GHK90, GM93]), le modèle de la double réservation inspiré d'Autolib' présente des interactions fortes qui disparaissent dans la limite champ-moyen.

Le processus limite : un processus de McKean–Vlasov

Je présente ici une approche heuristique où on "devine" le processus limite avant d'établir la convergence champ-moyen vers ce processus. Supposons que, pour $1 \leq i \leq N$, $(Z_i^N(t))$ converge en distribution vers un processus donné $(\overline{Z}(t)) = (\overline{R^r}(t), \overline{R}(t), \overline{V}(t), \overline{V^r}(t))$. Je me limiterai à deux termes qui apparaissent sur les deux premières équations d'évolution stochastiques (2.4).
Soit $\mathcal{P}_i^N$ la mesure aléatoire sur $\mathbb{R}_+$ définie par

$$\mathcal{P}_i^N([0,t]) = \int_0^t \sum_{j=1}^N \left(\sum_{l=1}^\infty \mathbf{1}_{\{l \leqslant X_{j,i}^N(s^-)\}} \mathcal{N}_{\nu,j,i,l}(ds) \right).$$

$\mathcal{P}_i^N$ est un processus ponctuel (dont la taille des sauts est 1) sur $\mathbb{R}_+$ avec comme processus croissant

$$\nu \int_0^t \sum_{j=1}^N X_{j,i}^N(s)\, ds = \nu \int_0^t R_i^{N,r}(s)\, ds.$$

Voir [Rob13] par exemple. En admettant la convergence en distribution du processus $(Z_i^N(t))$, on obtient que $\mathcal{P}_i^N$ converge vers un processus de Poisson inhomogène $\mathcal{P}^\infty$ d'intensité $(\nu \overline{R^r}(t))$ donné par

$$\mathcal{P}^\infty(dt) = \int_{\mathbb{R}_+} \mathbf{1}_{\{0 \leqslant h \leqslant \overline{R^r}(t^-)\}} \overline{\mathcal{N}}_{\nu,1}(dt, dh)$$

où $\overline{\mathcal{N}}_{\nu,1}$ est un processus de Poisson d'intensité $\nu dh dt$. De même, considérons aussi le processus ponctuel $\mathcal{Q}_i^N$ sur $\mathbb{R}_+$ défini par

$$\mathcal{Q}_i^N([0,t]) = \int_0^t \sum_{j=1}^N \mathbf{1}_{\{V_j^N(s^-)>0,\, S_i^N(s^-)<c\}}\, \mathcal{N}_{\lambda,j}^{U,N}(ds, \{i\})$$

dont le processus croissant est donné par

$$\lambda \int_0^t \mathbf{1}_{\{S_i^N(s)<c\}} \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{V_j^N(s)>0\}} ds.$$

Intuitivement, grâce à l'indépendance asymptotiques des stations, on obtient, quand N tend vers l'infini, un résultat de type Loi des Grands Nombres,

$$\frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{V_j^N(t)>0\}} \to \mathbb{P}(\overline{V}(t) > 0).$$

Ainsi, formellement, quand N tend vers l'infini, $\mathcal{Q}_i^N$ converge vers un processus de Poisson inhomogène $\mathcal{Q}^\infty$ d'intensité $(\lambda \mathbf{1}_{\{\overline{S}(t)<c\}} \mathbb{P}(\overline{V}(t)>0))$. En d'autres termes,

$$\mathcal{Q}^\infty(dt) = \int_{\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \mathbf{1}_{\overline{S}(t^-)<c\}}\mathbb{P}(\overline{V}(t^-)>0)\}} \overline{\mathcal{N}}_{\lambda,1}(dt,dh).$$

Ainsi, en passant à la limite quand N tend vers l'infini dans le système (2.4), on obtient formellement

$$\begin{cases} d\overline{R^r}(t) = & \int_{\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \mathbf{1}_{\{\overline{S}(t^-)<c\}}\mathbb{P}(\overline{V}(t^-)>0)\}} \overline{\mathcal{N}}_{\lambda,1}(dt,dh) - \int_{\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \overline{R^r}(t^-)\}} \overline{\mathcal{N}}_{\nu,1}(dt,dh), \\ d\overline{R}(t) = & \int_{\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \overline{R^r}(t^-)\}} \overline{\mathcal{N}}_{\nu,1}(dt,dh) - \sum_{l=1}^{+\infty} \mathbf{1}_{\{l\leqslant \overline{R}(t^-)\}} \mathcal{N}_{\mu,l}(dt), \\ d\overline{V}(t) = & \sum_{l=1}^{+\infty} \mathbf{1}_{\{l\leqslant \overline{R}(t^-)\}} \mathcal{N}_{\mu,l}(dt) - \int_{\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \mathbf{1}_{\{\overline{V}(t^-)>0\}}\mathbb{P}(\overline{S}(t^-)<K)\}} \overline{\mathcal{N}}_{\lambda,2}(dt,dh), \\ d\overline{V^r}(t) = & \int_{\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \mathbf{1}_{\{\overline{V}(t^-)>0\}}\mathbb{P}(\overline{S}(t^-)<c)\}} \overline{\mathcal{N}}_{\lambda,2}(dt,dh) - \int_{\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \overline{V^r}(t^-)\}} \overline{\mathcal{N}}_{\nu,2}(dt,dh). \end{cases} \tag{2.6}$$

Ces équations différentielles stochastiques dont les coefficients dépendent non seulement de l'état du processus inconnu, mais également de sa loi de probabilité, sont dites EDS de McKean–Vlasov, avec comme bruit dans notre cas des processus de Poisson. Également appelées EDS type champ-moyen, elles ont d'abord été étudiées en physique statistique par Kac [Kac56] et représentent en quelque sorte le comportement moyen d'un nombre infini de particules. Voir [McK66, Szn89] pour plus de détails. Un premier résultat de [FM25] établit l'existence et l'unicité du processus stochastique solution du système d'EDS (2.6). Soit $T>0$ fixé et soit $\mathcal{D}_T = \mathcal{D}([0,T],\mathcal{P}(\chi))$ l'ensemble des fonctions càd-làg de $[0,T]$ dans $\mathcal{P}(\chi)$ avec $\chi = \{(w,x,y,z)\in\mathbb{N}^4; w+x+y+z\leq c\}$.

Théorème 2.3.1 (*Processus de McKean–Vlasov*)

Pour tout $(w,x,y,z)\in\chi$, le système d'équations

$$\begin{cases} \overline{R^r}(t) = & w + \iint_{[0,t]\times\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \mathbf{1}_{\{\overline{S}(s^-)<c\}}\mathbb{P}(\overline{V}(s^-)>0)\}} \overline{\mathcal{N}}_{\lambda,1}(ds,dh) \\ & - \iint_{[0,t]\times\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \overline{R^r}(s^-)\}} \overline{\mathcal{N}}_{\nu,1}(ds,dh), \\ \overline{R}(t) = & x + \iint_{[0,t]\times\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \overline{R^r}(s^-)\}} \overline{\mathcal{N}}_{\nu,1}(ds,dh) \\ & - \int_0^t \sum_{l=1}^{+\infty} \mathbf{1}_{\{l\leqslant \overline{R}(s^-)\}} \mathcal{N}_{\mu,l}(ds), \\ \overline{V}(t) = & y + \int_0^t \sum_{l=1}^{+\infty} \mathbf{1}_{\{l\leqslant \overline{R}(s^-)\}} \mathcal{N}_{\mu,l}(ds) \\ & - \iint_{[0,t]\times\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \mathbf{1}_{\{\overline{V}(s^-)>0\}}\mathbb{P}(\overline{S}(s^-)<c)\}} \overline{\mathcal{N}}_{\lambda,2}(ds,dh) \\ \overline{V^r}(t) = & z + \iint_{[0,t]\times\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \mathbf{1}_{\{\overline{V}(s^-)>0\}}\mathbb{P}(\overline{S}(s^-)<K)\}} \overline{\mathcal{N}}_{\lambda,2}(ds,dh) \\ & - \iint_{[0,t]\times\mathbb{R}_+} \mathbf{1}_{\{0\leqslant h\leqslant \overline{V^r}(s^-)\}} \overline{\mathcal{N}}_{\nu,2}(ds,dh), \end{cases} \tag{2.7}$$

admet une solution unique $(\overline{R^r}(t), \overline{R}(t), \overline{V}(t), \overline{V^r}(t))$ *dans* $\mathcal{D}_T$.

Notons que la solution du système d'EDS (2.7) satisfait l'équation de Fokker–Planck (2.5).
L'étape suivante est de prouver la convergence champ-moyen pour le processus d'état simplifié $(Z^N(t))$, c'est-à-dire que la suite des processus de mesure empirique $(\Lambda^N(t))$ converge en distribution vers $(\overline{Z}(t))$. Cela signifie que, pour toute fonction f à support fini, la suite des processus $(\Lambda^N(t)(f))$ converge en distribution vers $(\mathbb{E}(f(Z(t))))$ avec, rappelons-le, le processus de mesure empirique $\Lambda^N(t)$ de $(Z_i^N(t), 1 \leqslant i \leqslant N)$ défini pour toute fonction f à support fini dans χ, par

$$\Lambda^N(t)(f) = \frac{1}{N}\sum_{i=1}^{N} f(Z_i^N(t)) = \frac{1}{N}\sum_{i=1}^{N} f(R_i^{r,N}(t), R_i^N(t), V_i^N(t), V_i^{r,N}(t)).$$

Comme le processus $(Z^N(t))$ n'est pas markovien, le processus de mesure empirique $(\Lambda^N(t))$ ne l'est pas.

Théorème 2.3.2 (*Convergence champ-moyen*)

La suite des processus de mesure empirique $(\Lambda^N(t))$ *converge en distribution vers un processus* $(\Lambda(t)) \in \mathcal{D}([0,T], \mathcal{P}(\chi))$ *défini par, pour* f *à support fini dans* χ*,*

$$\Lambda(t)(f) = \mathbb{E}(f(\overline{Z}(t)))$$

avec $(\overline{Z}(t))$ *l'unique solution de l'équation* (2.7)*. De plus, pour tout* $k \geqslant 1$ *et pour* $1 \leqslant i_1 < \ldots < i_k \leqslant N$*, la suite à support fini des marginales* $(Z_{i_1}^N(t), \ldots, Z_{i_k}^N(t))$ *converge en distribution vers* $(\overline{Z}_{i_1}(t), \ldots, \overline{Z}_{i_k}(t))$*, où* $(\overline{Z}_{i_1}(t))$*,...,* $(\overline{Z}_{i_k}(t))$ *sont des variables aléatoires indépendantes de même distribution que* $(\overline{Z}(t))$.

La dernière propriété est dite propagation du chaos. Voir [Szn89] pour plus de détails.

Le point d'équilibre du processus limite

L'étude du comportement du système en temps long nécessite de poser le problème de l'existence et surtout de l'unicité de la mesure invariante du processus limite $(\overline{Z}(t))$, puisque, pour les processus de Markov inhomogènes, il peut y avoir plusieurs mesures invariantes. Dans [FM25], nous prouvons l'existence d'une mesure invariante unique dans un cadre restreint. La preuve est basée sur trois arguments principaux. Tout d'abord, en appliquant la théorie des files d'attente, le processus limite de McKean-Vlasov $(\overline{Z}(t))$ est identifié à un tandem de quatre files d'attente à capacité globale fixée c et avec une mesure invariante de forme-produit explicite bien connue. Cela signifie que $\overline{R^r}(t)$, $\overline{R}(t)$, $\overline{V}(t)$ et $\overline{V^r}(t)$ sont respectivement les nombres de clients dans

- la première file d'attente, une $M/M/\infty$ avec un taux de service de ν,
- la seconde, une $M/M/\infty$ avec un taux de service de μ,
- la troisième, une $M/M/1$ avec un taux de service variable $\lambda\mathbb{P}(\overline{S}(t) < c)$ à l'instant t.
- et la dernière, une $M/M/\infty$ avec un taux de service ν,

tandis que le processus d'arrivée est un processus de Poisson inhomogène d'intensité $\lambda\mathbb{P}(\overline{V}(t) > 0)dt$. Voir la Figure 2.6 pour une illustration. Soient η_1, ρ_1, ρ_2 et η_2 les rapports entre le taux d'arrivée et le taux de service pour les quatre files d'attente de gauche à droite (comme indiqué sur la Figure 2.6). Par définition, on obtient

$$\begin{cases} \eta_1(t) &= \dfrac{\lambda}{\nu}\mathbb{P}(\overline{V}(t) > 0), \quad \rho_1(t) = \dfrac{\lambda}{\mu}\mathbb{P}(\overline{V}(t) > 0), \\ \rho_2(t) &= \dfrac{\mathbb{P}(\overline{V}(t) > 0)}{\mathbb{P}(\overline{S}(t) < c)}, \quad \eta_2(t) = \dfrac{\lambda}{\nu}\mathbb{P}(\overline{V}(t) > 0), \end{cases} \tag{2.8}$$

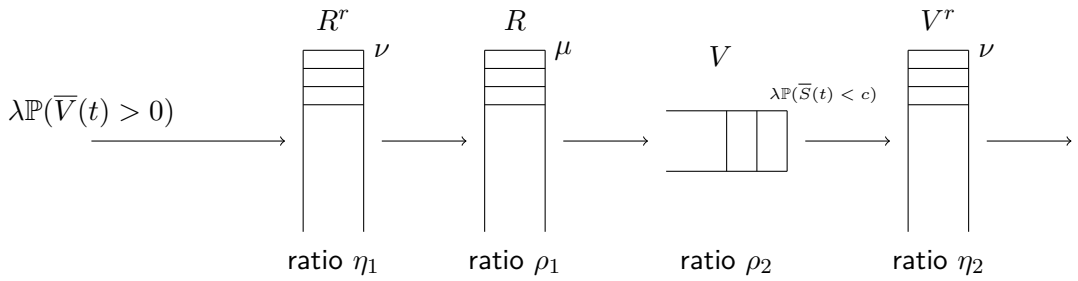

Figure 2.6 – $(\overline{Z}(t))$ vu comme un tandem de 4 files d'attente. Les files verticales sont des $M/M/\infty$, l'horizontale une $M/M/1$. La capacité globale est c.

ce qui nous permet d'écrire la mesure invariante du quadruplet $(\overline{R^r}(t), \overline{R}(t), \overline{V}(t), \overline{V^r}(t))$ comme suit

$$\pi_{j,k,l,m}(\rho) = \frac{1}{Z(\rho)} \frac{\eta_1^j \rho_1^k}{j!k!} \rho_2^l \frac{\eta_2^m}{m!} \tag{2.9}$$

où, pour alléger les notations, $(\eta_1, \rho_1, \rho_2, \eta_2)$ est noté ρ et la constante de normalisation est donnée par

$$Z(\rho) = \sum_{j+k+l+m \leqslant c} \frac{\eta_1^j \rho_1^k}{j!k!} \rho_2^l \frac{\eta_2^m}{m!}.$$

Si le tandem des 4 files d'attente décrit ci-dessus est à l'équilibre, alors

$$0 = \pi(\rho(t)) L_{\rho(t)}$$

où $L_{\rho(t)}$ désigne le générateur infinitésimal du processus $(\overline{Z}(t))$ et $\rho(t) = (\eta_1, \rho_1, \rho_2, \eta_2)(t)$ est donné par l'équation (2.8). Le point d'équilibre est alors la mesure de probabilité $\pi(\rho)$ définie par (2.9) telle que $\rho = (\eta_1, \rho_1, \rho_2, \eta_2)$ satisfait le système suivant

$$\begin{cases} \eta_1 &= \dfrac{\lambda}{\nu}(1 - \pi_{0V}(\rho)), \\ \rho_1 &= \dfrac{\lambda}{\mu}(1 - \pi_{0V}(\rho)), \\ \rho_2 &= \dfrac{1 - \pi_{0V}(\rho)}{1 - \pi_S(\rho)}, \\ \eta_2 &= \dfrac{\lambda}{\nu}(1 - \pi_{0V}(\rho)) \end{cases} \tag{2.10}$$

avec $\pi_{0V}(\rho) = \sum_{j+l+m \leqslant c} \pi_{j,0,l,m}(\rho)$ et $\pi_S(\rho) = \sum_{j+k+l+m=c} \pi_{j,k,l,m}(\rho)$. Considérant ce tandem de files d'attente comme une station, $\pi_{0V}(\rho)$ est la probabilité qu'aucune voiture ne soit disponible dans cette station et $\pi_S(\rho)$ celle que la station soit saturée, c'est-à-dire qu'aucune place de parking n'y soit disponible.
Enfin, rappelons que le nombre moyen de voitures par station est fixé à s, on obtient alors une contrainte qui vient s'ajouter au système (2.10)

$$s = \sum_{j+k+l+m \leqslant c} (k + l + m) \pi_{j,k,l,m}(\rho).$$

Après simplification du système (2.10), le problème de l'existence et de l'unicité de la mesure invariante du processus de Markov inhomogène $(\overline{Z}(t))$ est réduit à un problème de point fixe en dimension 2. Le théorème d'inversion globale et une propriété de monotonie nous permettent de conclure. Les deux derniers arguments sont basés sur des calculs combinatoires. On obtient alors le théorème suivant

Théorème 2.3.3

Pour ν assez grand, il existe un unique point d'équilibre pour l'équation de Fokker–Planck (2.5) *donné par la mesure de probabilité $\pi(\rho)$ définie par l'équation* (2.9) *où $\rho = (\frac{\mu}{\nu}\rho_1, \rho_1, \rho_2, \frac{\mu}{\nu}\rho_1)$ et (ρ_1, ρ_2) désigne l'unique solution de*

$$\begin{cases} \rho_1 &= \frac{\lambda}{\mu}(1 - \pi_{0V}(\rho)), \\ s &= \sum_{j+k+l+m \leqslant c}(k+l+m)\pi_{j,k,l,m}(\rho). \end{cases} \tag{2.11}$$

La monotonie est simplement prouvée lorsque le temps de réservation moyen $1/\nu$ est suffisamment petit. Nous sommes convaincus que cette hypothèse est technique mais la preuve nécessite d'autres outils. Le système (2.11) permet d'obtenir numériquement (ρ_1, ρ_2) pour un paramètre s fixé puisque les probabilités $\pi_{0V}(\rho)$ et $\pi_{j,k,l,m}(\rho)$ ne sont que des fractions rationnelles de ρ_1 et ρ_2. Ceci nous permet de mesurer la performance du système en gardant la même métrique que les autres modèles, à savoir la proportion limite de stations sans voiture ou place de parking disponible.

2.4 Pour aller plus loin : la politique de réservation

L'opérateur d'autopartage Communauto propose deux types de formules à Montréal : avec et sans stations physiques. On se focalise ici sur l'offre basée sur des stations physiques où les usagers peuvent utiliser une voiture garée dans l'une des stations réparties dans la ville afin d'effectuer un trajet puis de ramener la voiture à la station de départ. Ce sont des trajets en *boucle*. Au départ d'une voiture, la place de parking lui reste dédiée, donc le problème de saturation ne se pose pas pour l'usager. Dans chaque station, il y a autant de voitures que de places de parking. Ainsi, les stations interagissent entre elles seulement à travers le déport d'usagers cherchant une voiture disponible.

Cette formule est principalement destinée aux plus long trajets, planifiés à l'avance, tels que des déplacements pour une longue période (week-end, vacances...). En pratique, les usagers ont accès à une application pour réserver une voiture. Le moment où un usager réserve un créneau futur est appelé *temps d'inscription*. La réservation peut s'effectuer jusqu'à un mois à l'avance. On peut séparer en deux le temps entre l'inscription d'une réservation et la fin de la réservation : Le temps s'écoulant entre l'inscription et le début de la réservation est appelé *temps de pré-réservation*, le *temps de réservation* correspond lui à la durée de réservation elle-même (et non durée de trajet).

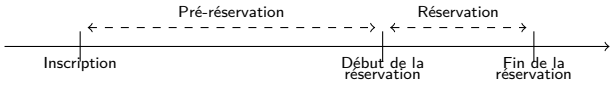

Figure 2.7 – Schéma d'une réservation

Précisons qu'il est possible pour un usager d'annuler une réservation effectuée ou encore de rendre la voiture avant la fin prévue de la réservation. Dans ce dernier cas, on parlera de *fin de réservation anticipée*. Une étape préliminaire d'analyse de données opérateur sur l'année 2021 a permis de comprendre le fonctionnement du système, de visualiser le caractère stochastique des réservations (on observe ainsi sur la Figure 2.8 de grandes fluctuations) et d'approcher les lois des différentes durées de temps en question.

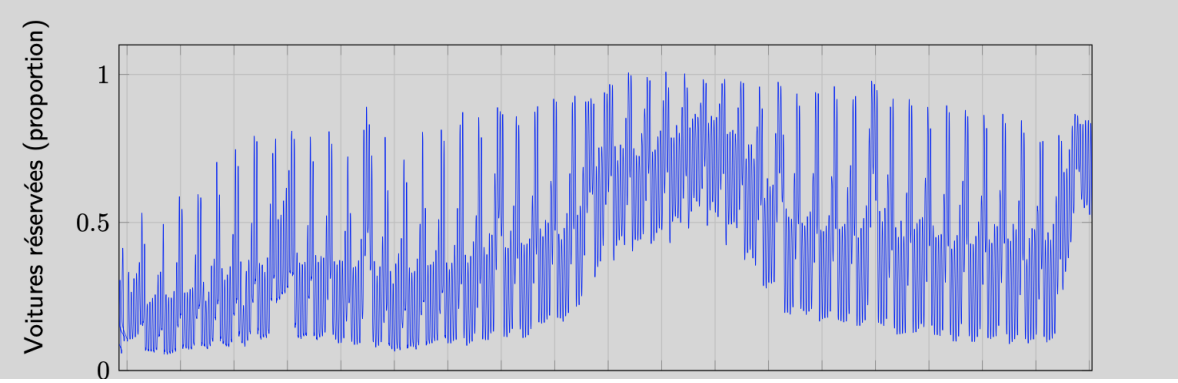

Figure 2.8 – Evolution des réservations au cours de l'année 2021

Hypothèses de modélisation

Précisons qu'on considère que les voitures sont indiscernables et que les usagers ne réservent pas une voiture spécifique mais une voiture quelconque du système. Lorsque la demande est forte, on fait l'hypothèse qu'un usager est près à se déplacer "à l'autre bout de la ville" pour une voiture disponible pour le créneau horaire recherché. Autrement dit, l'usager réserve un créneau et non une voiture à une station donnée. On désigne par C le nombre total de voitures du systèmes à travers l'ensemble des stations. Ce nombre sera appelé *capacité du système*, à distinguer de la notion de capacité d'une station.
Ainsi, nous faisons le choix d'un modèle stochastique dans lequel C voitures sont disponibles pour des demandes de réservation qui arrivent au taux λ. Lorsque l'usager effectue une demande de réservation, si une voiture est disponible pour la totalité du créneau horaire futur recherché, une réservation est effectuée pour ce créneau. Dans le cas contraire, la demande est rejetée.

Capacité infinie : un premier modèle simplifié

Dans ce cas purement théorique, toutes les demandes de réservation sont acceptées puisque la capacité (nombre total de voitures) est infinie.
On considère pour l'instant un modèle simplifié où les usagers réservent à l'avance un créneau dans le futur et ne peuvent ni l'annuler ni le finir prématurément. Les réservations sont modélisées par un processus de Poisson marqué (t_n, p_n, σ_n) sur $\mathbb{R}_+^3$ avec :

- t_n est l'instant d'inscription de la réservation, suivant un processus de Poisson de paramètre λ.
- p_n est la durée de pré-réservation de la réservation.
- σ_n est la durée de réservation.

La suite (p_n, σ_n) est i.i.d et (t_n) et (p_n, σ_n) sont indépendantes. On note (p, σ) un couple de variables aléatoires ayant même loi que les (p_n, σ_n).
On introduit aussi la notion de *nombre de réservations actives* à un instant donné correspondant au nombre de véhicules réservés à cet instant, ayant fini la période de pré-réservation mais pas encore celle de réservation proprement dite. Il vaut

$$L(t) = \sum_{n=0}^{\infty} \mathbf{1}_{\{t_n+p_n \leqslant t < t_n+p_n+\sigma_n\}}. \tag{2.12}$$

En notant $B_t = \{(u, x, y) \in (\mathbb{R}^+)^3 \,|u + x < t < u + x + y\}$, on obtient que $L(t) = \mathcal{N}(\mathbf{1}_{B_t})$. Et donc, on peut montrer que, $\forall t > 0$, on a

$$\mathbb{E}(L(t)) = \mathbb{E}\big(\min((t-p)^+, \sigma)\big).$$

Dans le cas d'une capacité infinie, si on fait abstraction de la pré-réservation ($p_n \equiv 0$), notre modèle est simplement une file d'attente $M/G/\infty$. Le nombre de client dans une telle file est donnée par

$$L(t) = \sum_{n=0}^{\infty} \mathbf{1}_{\{t_n \leqslant t < t_n + \sigma_n\}} = \mathcal{N}(\mathbf{1}_{A_t})$$

avec $A_t = \{(u,x) \in \left(\mathbb{R}^+\right)^2 | u < t < u + x\}$. Afin de décrire le comportement stationnaire d'une telle file, l'étude de la limite fluide du processus $(L(t))$ s'avère être très adaptée. En effet, cette méthode de renormalisation est l'une des rares techniques disponibles pour les cas où les services des clients sont généraux et non plus exponentiels. On définit le processus normalisé $\overline{L}_N(t) := \dfrac{L_N(t)}{N}$ où $L_N(t)$ est un processus similaire à $(L(t))$ avec un taux d'arrivée $N\lambda$. Ceci revient, en un certain sens, à accélérer le temps et renormaliser l'espace par un même facteur d'échelle N. La limite, quand N tend vers l'infini, de $(\overline{L}_N(t))$ est appelée *limite fluide* du processus d'origine $(L(t))$. En suivant la même démarche que dans [FJ07], on décompose le terme non martingale dans (2.12)

$$\begin{aligned} L_N(t) = \mathcal{N}_N(\mathbf{1}_{B_t}) &= \int_{\mathbb{R}^3_+} \mathbf{1}_{B_t}(u,x,y)\mathcal{N}_N(du,dx,dy) \\ &= \int_{\mathbb{R}^3_+} \mathbf{1}_{\{0\leq u+x\leq t)\}}\mathcal{N}_N(du,dx,dy) - \int_{\mathbb{R}^3_+} \mathbf{1}_{\{0\leq u+x+y\leq t)\}}\mathcal{N}_N(du,dx,dy) \\ &= \int_{\mathbb{R}^3_+} \mathbf{1}_{\{0\leq u\leq t)\}}\mathcal{N}_{N,\phi_1}(du,dx,dy) - \int_{\mathbb{R}^3_+} \mathbf{1}_{\{0\leq u\leq t)\}}\mathcal{N}_{N,\phi_2}(du,dx,dy) \end{aligned}$$

où

- $\mathcal{N}_N$ est un processus de Poisson de paramètre $N\lambda$,
- $\mathcal{N}_{N,\phi_i}$ est un processus de Poisson dont l'intensité est l'image de l'intensité de $\mathcal{N}_N$ par ϕ_i
- $\phi_1 : (u,x,y) \mapsto (u+x,x,y)$ et $\phi_2 : (u,x,y) \mapsto (u+x+y,x,y)$

de façon à décomposer le processus normalisé $\overline{L}_N(t)$ en un terme martingale et un autre déterministe

$$\overline{L}_N(t) = \frac{L_N(t)}{N} = \frac{M_{N,1}(t)}{N} - \frac{M_{N,2}(t)}{N} + \lambda\,\mathbb{E}\big(\min((t-p)^+,\sigma)\big).$$

où $(M_{N,i}(t))$ sont deux martingales, chacune par rapport à sa filtration. Ceci n'empêche pas de montrer que chacun des deux termes martingales normalisées tend vers 0 quand N devient grand grâces aux inégalités de Cauchy-Schwarz et de Doob.
La même démarche est vraie quand on tient compte de la possibilité d'annulation et de retour anticipé. Notons alors par

- a_n le temps d'annulation de la réservation (le temps entre l'inscription et l'annulation de la réservation),
- r_n le temps de retour anticipé du véhicule.

On peut alors énoncer le résultat suivant

Théorème 2.4.1 (*Limite fluide pour une capacité infinie*)

Le processus normalisé $(\overline{L}_N(t))$ converge quand N tend vers l'infini vers

$$x(t) = \lambda\,\mathbb{E}\big(\mathbf{1}_{\{p\leq a\}}\min((t-p)^+,\sigma,r)\big).$$

Ainsi, le comportement macroscopique stationnaire du processus initial $(L(t))$ est donné par

$$\lim_{t\to+\infty} x(t) = \lambda\,\mathbb{E}\big(\mathbf{1}_{\{p\leq a\}}\min(\sigma,r)\big).$$

La Figure 2.9 est une simulation numérique avec de données réelles qui valide bien le fait que la limite fluide obtenue ici est un bon indicateur du comportement du système, avec ou sans pré-réservation.

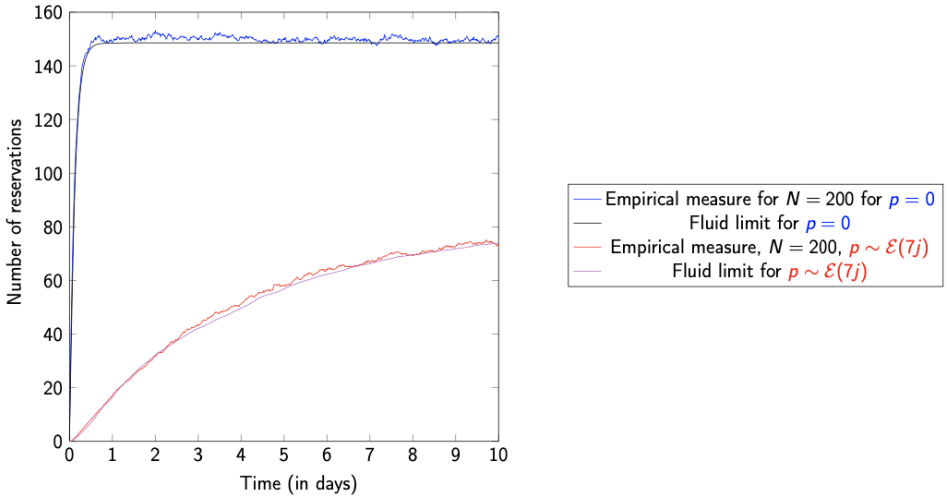

Figure 2.9 – Comparaison de la limite fluide et de la mesure empirique.

Capacité finie : le modèle "Tetris"

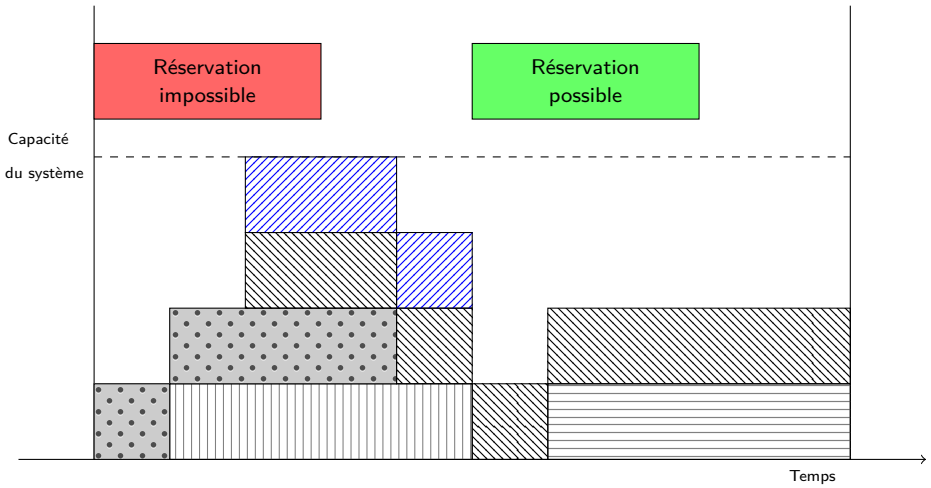

Figure 2.10 – Schéma représentant le modèle "Tetris"

Pour un modèle plus réaliste, il faut tenir compte de la capacité. Afin de visualiser la condition d'acceptabilité des réservations, on peut représenter les réservations successives comme des pièces d'une certaine longueur proportionnelle à la longueur du créneau réservé. On peut alors représenter les réservations successives comme différentes pièces tombant dans une grille de "Tetris" infinie, l'axe des abscisses représentant l'axe temporel. Chaque réservation descend alors sur le créneau que le client souhaite réserver et la réservation est acceptée uniquement si ce bloc peut se retrouver en totalité en dessous de la capacité du système. La Figure 2.10 schématise ce modèle de déposition pour un système d'une capacité maximale égale $C = 4$.
Pour le moment, seule une bande supérieure pour la limite fluide $x_C(t)$ du processus $(L(t))$ est établie.

$$x_C(t) \leq \min\Big(\underbrace{\lambda\,\mathbb{E}(\mathbf{1}_{\{p\leq a\}}\min((t-p)^+,\sigma,r))}_{x(t)},C\Big) \tag{2.13}$$

où $x(t)$ est la limite fluide obtenue pour une capacité infinie. Dans le cas d'une capacité finie, tout reste à faire, notamment la question d'une limite fluide explicite ou encore l'existence d'une transition de phase, mise en évidence par simulations, entre deux régimes, un saturé et un non saturé. En effet, dans le régime saturé, on observe l'apparition d'une valeur à long terme de la limite fluide, autrement dit un nombre de réservation maximal, $C^* < C$. Ceci n'est pas le cas pour le régime non saturé. La Figure 2.11 est une comparaison entre le comportement de la limite fluide "réelle" du système à travers sa mesure empirique et celle à laquelle on s'attend intuitivement, soit la borne supérieure (2.13). On y voit apparaître deux régimes. Ces questions font l'objet de travaux en cours. Remarquons aussi que ces questions autour de la réservation dans le futur dépassent le cadre des systèmes de partages de véhicules et sont valables pour de nombreux domaines (hôtels, gestion de salles à usage collectif . . .)

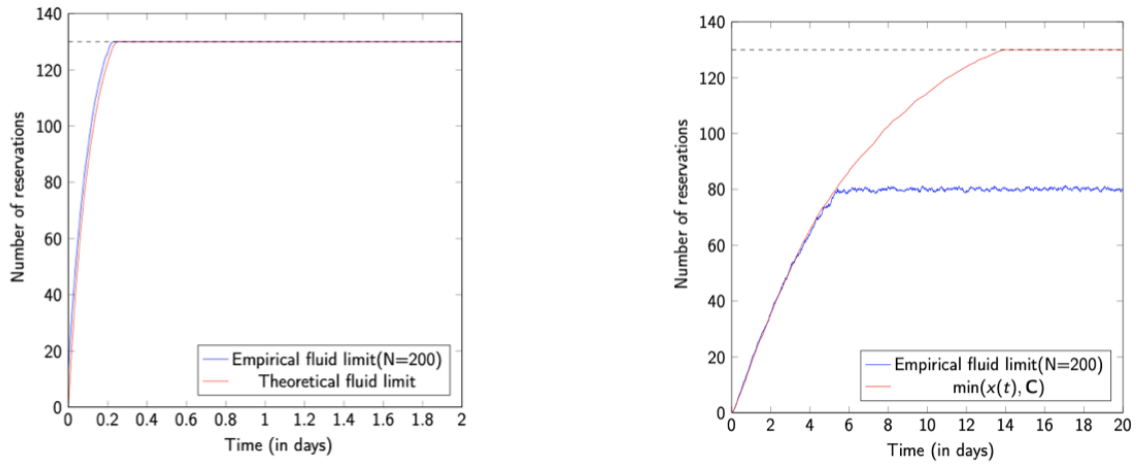

Figure 2.11 – Comparaison des limites fluides empirique et intuitive. Mise en évidence de deux régimes.

Chapitre 3

Le free-floating : un environnement qui varie rapidement

Collaborations avec Christine Fricker, Alessia Rigonat et Martin Trépanier.

3.1 Motivation

Dans le chapitre précédant, nous avons évoqué brièvement les systèmes de partages de voitures sans stations physiques, appelés *free-floating*. Dans la littérature transport, ces systèmes sont assimilés à des systèmes d'autopartage classiques avec des pseudo-stations (par découpage de la zone de service en un maillage composé de petites zones géographiques dont la superficie varie entre $0.25\,km^2$ et $1\,km^2$). Reste alors le paramètre capacité à définir. Dans cette première approche, ces pseudo-stations sont considérées à capacité fixe. En pratique, l'opérateur considère le nombre maximum de voitures partagées garées simultanément dans la pseudo-station comme étant sa capacité. Ce choix n'est pas pertinent. En effet, les voitures *free-floating* et les voitures privées se partagent les places de parking dans ces pseudo-stations. Les voitures *free-floating* étant beaucoup moins nombreuses que les voitures particulières, cette capacité globale est donc principalement occupée par ces dernières. Ainsi, les voitures *free-floating* disposent d'une capacité résiduelle aléatoire. Cette caractéristique est spécifique aux systèmes d'autopartage en *free-floating*. Il s'agit d'une différence majeure par rapport aux systèmes basés sur des stations physiques, où la capacité d'une station est fixe.

3.2 Le modèle

Dans [FMRT25, FMR25], nous avons proposé un nouveau modèle stochastique, dédié au *free-floating* prenant en compte les interactions entre les voitures privées et les voitures partagées à travers le partage des places de parking sur l'espace publique. Ce point de vue est original et semble beaucoup plus pertinent pour le système que l'approche classique utilisée jusqu'à présent. Considérons un système d'autopartage en *free-floating* de grande taille, c'est-à-dire avec un grand nombre de pseudo-stations noté N. Ainsi, N sera notre paramètre d'échelle. Nous partons du principe que la capacité totale de parking d'une pseudo-stations de superficie $1\,km^2$ est grande, disons de l'ordre de N, ainsi que la taille de la flotte, c'est à dire le nombre des voitures partagées, $M \sim s\,N$ où s est le paramètre de dimensionnement recherché par l'opérateur.
L'homogénéité étant naturelle pour une approche champ-moyen, nous considérons un ensemble de N nœuds (pseudo-stations), chacun ayant une capacité fixe cN, où c est une constante positive. Il y a deux types de clients (voitures). Un nombre fixe $M \sim s\,N$ de clients de type 1 (les voitures

partagées) se déplacent parmi les N nœuds. Au taux λ, un de ces clients quitte un nœud. Après un temps de trajet de moyenne $1/\mu$, il rejoint un autre nœud choisi au hasard. Si le nœud choisi est saturé, c'est-à-dire qu'il n'y a pas de place disponible, un second trajet de même loi est entamé vers un autre nœud choisi au hasard et ainsi de suite jusqu'à ce qu'à trouver un nœud non saturé. Les clients de type 2 (les voitures privées) entrent et sortent du système, sans se déplacer entre ses nœuds. En effet, ils arrivent à un nœud donné à un taux αN et y restent pendant un temps de service de moyenne $1/\beta$ avant de quitter le système. S'il n'y a pas de place disponible à ce nœud, le client de type 2 quitte le système. Intuitivement, le choix d'un tel taux αN implique que le nombre de clients de type 2 à un nœud donné est d'ordre N. En revanche, le nombre de clients de type 1 à un nœud donné est d'ordre M/N soit borné. Ceci reflète le déséquilibre numérique entre les voitures privées et les voitures partagées dans les systèmes réels. Ce choix de modélisation est conforté par l'analyse de données de l'opérateur Communauto à Montréal où on comptait en 2023 environ 850 voitures partagées contre $800\,000$ voitures privées pour 150 pseudo-stations de $1\,km^2$ chacune.
Tous les temps d'inter-arrivée, de service et de trajet sont indépendants avec des distributions exponentielles.

3.3 Analyse multi-échelle et approche champ-moyen

Dans notre modèle, sans les clients de type 2, le système se réduit à un réseau fermé de Jackson où les M clients se déplacent entre N nœuds se comportant en tant que des files d'attente à un serveur et à capacité cN alors que les routes se comportent comme des files d'attente à une infinité de serveurs. Sans les clients de type 1, chaque nœud peut être considéré comme une file d'attente $M/M/cN/cN$ avec un taux d'arrivée accéléré. Cependant, la dynamique de ces files d'attente est perturbée par la présence d'un petit nombre de clients de type 1 qui, en se déplaçant entre les nœuds, créent l'interaction dans le système. Intuitivement, la perturbation qu'ils créent est marginale en raison de la différence d'ordre de grandeur du nombre de clients de chaque type. Ainsi, le comportement d'un nœud donné est celui d'une file d'attente $M/M/cN/cN$ bien connue, voir [Rob13] par exemple. On obtient alors deux processus évoluant à deux vitesses différentes. Le nombre d'emplacements disponibles est le processus rapide alors que le nombre de clients de cette file d'attente $M/M/cN/cN$, renormalisé par N, est le processus lent. En général, le processus lent considère le processus rapide comme un environnement qui varie rapidement. Voir par exemple [BMP10]. Cette file d'attente et plus généralement les grands réseaux avec perte [HK94] présentent un principe de moyennisation stochastique *stochastic averaging* où le processus lent "voit" le processus rapide à la stationnarité. Par exemple, pour les réseaux avec perte, la probabilité de blocage du processus lent dans le régime de forte charge est une quantité moyenne du processus rapide du nombre d'emplacements vides. Des résultats de convergence sont obtenus pour le processus de mesure d'occupation et différentes échelles de temps sont souvent utilisées pour étudier le comportement du modèle. Voir [FRZ23] pour une illustration de ces techniques.

De plus, il existe une transition de phase pour la file d'attente $M/M/cN/cN$ entre un régime sous-critique (à faible charge) et un régime sur-critique (à forte charge). Voir par exemple [Rob13].
Nous avons l'intuition que notre modèle présente le même phénomène de moyennisation stochastique et de transition de phase. La transition de phase ne devrait dépendre que des paramètres de l'environnement, c'est-à-dire de la dynamique des clients de type 2 et de la capacité c.
Dans notre cas, un défi supplémentaire vient s'ajouter suite à la présence de N nœuds, ce qui implique des résultats de convergence champ-moyen. En effet, l'approche classique champ-moyen ne s'applique pas aux processus à plusieurs échelles de temps. Peu d'articles traitent de la limite champ-moyen pour les dynamiques multi-échelles. Voir par exemple [BMP10, CBLW21, CBM+21].

Processus d'état

Notons par $m_i^N(t)$ le nombre d'emplacements disponibles, $V_i^N(t)$ le nombre de clients de type 1 et $X_i^N(t)$ le nombre de clients de type 2 à un nœud donné i pour $i = 1, ..., N$ au temps t. Rappelons, pour $i = 1 \dots N$ et pour $t \geqslant 0$, la relation suivante

$$m_i^N(t) + V_i^N(t) + X_i^N(t) = cN.$$

Le processus $(m^N(t), V^N(t)) = ((m_i^N(t), V_i^N(t)),\ 1 \leqslant i \leqslant N)$ est un processus de Markov sur l'espace d'état

$$\mathcal{S}^N = \left\{ (m_i, v_i)_{1 \leqslant i \leqslant N} \in \mathbb{N}^{2N}, m_i + v_i \leqslant cN,\ \sum_{k=1}^{N} v_k \leqslant M_N \right\}$$

avec des transitions qui modifient l'état d'un nœud à la fois. Par exemple, pour le nœud i, on a

$$(m_i, v_i) \to \begin{cases} (m_i + 1, v_i) & \text{au taux } \beta(cN - m_i - v_i) \\ (m_i - 1, v_i + 1) & \text{au taux } \mathbf{1}_{\{m_i > 0\}} \mu (M_N - \sum_{l=1}^{N} v_l)/N \\ (m_i + 1, v_i - 1) & \text{au taux } \lambda \mathbf{1}_{\{v_i > 0\}} \\ (m_i - 1, v_i) & \text{au taux } \alpha N \mathbf{1}_{\{m_i > 0\}}. \end{cases}$$

Equations différentielles stochastiques

Le processus $((m_i^N(t), V_i^N(t)), 1 \leqslant i \leqslant N)$ peut être vu comme solution des équations différentielles stochastiques suivantes

$$\begin{aligned} dm_i^N(t) =& \mathbf{1}_{\{V_i^N(t^-) > 0\}} \mathcal{N}_\lambda^i(dt) + \sum_{j=1}^{\infty} \mathbf{1}_{\{j \leqslant CN - m_i^N(t^-) - V_i^N(t^-)\}} \mathcal{N}_\beta^{i,j}(dt) \\ & - \left(\sum_{j=1}^{\infty} \mathbf{1}_{\{j \leqslant M - \sum_{l=1}^{N} V_l^N(t^-), m_i^N(t^-) > 0\}} \bar{\mathcal{N}}_\mu^j(dt, \{i\}) + \mathbf{1}_{\{m_i^N(t^-) > 0\}} \mathcal{N}_{\alpha N}^i(dt) \right) \\ dV_i^N(t) =& \sum_{j=1}^{\infty} \mathbf{1}_{\{j \leqslant M - \sum_{l=1}^{N} V_l^N(t^-), m_i^N(t^-) > 0\}} \bar{\mathcal{N}}_\mu^j(dt, \{i\}) - \mathbf{1}_{\{V_i^N(t^-) > 0\}} \mathcal{N}_\lambda^i(dt). \end{aligned}$$

où $(\mathcal{N}_\lambda^i)$, $(\mathcal{N}_{\alpha N}^i)$ et $(\mathcal{N}_\beta^{i,j})$ sont des suites de processus de Poisson indépendants sur $\mathbb{R}_+$ avec les intensités respectives λdt, $\alpha N dt$ et βdt, $\bar{\mathcal{N}}_\mu^j$ est une suite de processus de Poisson indépendants sur $\mathbb{R}_+^2$ d'intensité $\mu dt \otimes \mathcal{U}$, avec $\mathcal{U}$ la distribution uniforme sur $\{1, ..., N\}$. Ainsi, $\bar{\mathcal{N}}_\mu^j(\cdot, \{i\})$ est un processus de Poisson sur $\mathbb{R}_+$ d'intensité $\mu dt/N$.

Processus de mesure empirique

Considérons le processus mesure empirique $(\Lambda^N(t))$ associé à $(m_i^N(t), V_i^N(t), 1 \leqslant i \leqslant N)$, pour f fonction à support fini sur $\mathbb{N}^2$ et $t \geqslant 0$, défini par

$$\Lambda^N(t)(f) = \frac{1}{N} \sum_{i=1}^{N} f(m_i^N(t), V_i^N(t)).$$

La première étape est d'écrire les équations d'évolution du processus $(f(m_i^N(t), V_i^N(t)),\ 1 \leqslant i \leqslant N)$. Il serait alors utile d'introduire les notations suivantes, pour (m, v) et $(i, j) \in \mathbb{N}^2$,

$$\Delta_{i,j}(f)(m, v) = f(m + i, v + j) - f(m, v) \tag{3.1}$$

et d'opter pour une réécriture intégrale du processus de mesure empirique $\Lambda^N(t)(f)$ comme suit

$$\Lambda^N(t)(f) = \frac{1}{N}\sum_{i=1}^{N} f(m_i^N(t), V_i^N(t)) = \int_{\mathbb{N}^2} f(m,v)\Lambda^N(t)(dm,dv)$$

à l'aide de la mesure de probabilités sur $\mathbb{N}^2$

$$\Lambda^N(t) = \frac{1}{N}\sum_{i=1}^{N} \delta_{(m_i^N(t), V_i^N(t))}.$$

On obtient alors l'équation d'évolution fonctionnelle pour le processus de mesure empirique

$$\begin{aligned}
&\Lambda^N(t)(f) = \Lambda^N(0)(f) \qquad (3.2)\\
&+ \lambda \int_0^t \int_{\mathbb{N}^2} \Delta_{(1,-1)}(f)(m,v)\mathbf{1}_{\{v>0\}}\Lambda^N(s)(dm,dv)ds\\
&+ \mu \int_0^t \left(\frac{M}{N} - \int_{\mathbb{N}^2} v\Lambda^N(s)(dm,dv)\right) \int_{\mathbb{N}^2} \Delta_{(-1,1)}(f)(m,v)\mathbf{1}_{\{m>0\}}\Lambda^N(s)(dm,dv)ds\\
&+ \beta cN \int_0^t \int_{\mathbb{N}^2} \Delta_{(1,0)}(f))(m,v)\Lambda^N(s)(dm,dv)ds - \beta \int_0^t \int_{\mathbb{N}^2} (m+v)\Delta_{(1,0)}(f)(m,v)\Lambda^N(s)(dm,dv)ds\\
&+ \alpha N \int_0^t \int_{\mathbb{N}^2} \Delta_{(-1,0)}(f)(m,v)\mathbf{1}_{\{m>0\}})\Lambda^N(s)(dm,dv)ds + \frac{1}{N}\sum_{i=1}^{N} M_{i,f}^N(t)
\end{aligned}$$

où, pour $i = 1, ..., N$, le terme martingale $(M_{i,f}^N(t))$, de carré intégrable, est explicite.

Deux régimes, deux échelles de temps

Notre objectif est d'étudier le comportement du processus d'état à l'aide de l'équation d'évolution du processus de mesure empirique fonctionnelle. Dans l'équation (3.2), certains termes contiennent l'indicatrice du processus rapide $(m^N(t))$. Intuitivement, cette indicatrice $\mathbf{1}_{\{m_i^N(t)>0\}}$ lorsque N devient grand se comporte comme la probabilité stationnaire que le nombre d'emplacements disponibles soit différent de 0. C'est le principe de la moyennisation stochastique. Il existe donc deux régimes en fonction de cette probabilité : lorsqu'elle est égale à 1, on parle de régime sur-critique et, dans le cas contraire, de régime sous-critique.

— *Régime sur-critique*. Quand la charge $\alpha N/\beta$ dépasse la capacité cN, ou encore $\alpha/\beta > c$, le système est surchargé et il y a alors quelques emplacements disponibles.

— *Régime sous-critique*. Si $\alpha/\beta < c$, le système est sous-chargé et le nombre d'emplacements disponibles est alors autour de $(c - \alpha\ /\beta)N$.

Nous nous focalisons sur le régime sur-critique, plus intéressant techniquement mais aussi d'un point de vue applicatif (*l'autopartage aurait-il un impact négatif sur la congestion en cas de fort trafic ?*). Les processus $(m^N(t))$ et $(V^N(t))$ n'évoluent pas à la même vitesse en raison de la différence des ordres de grandeur de leurs taux respectifs. Le processus $(m^N(t))$ est rapide puisque ses taux sont d'ordre N alors que le processus $(V^N(t))$ est lent avec des taux d'ordre 1. Pour saisir le comportement stationnaire du processus lent $(V^N(t))$, le système est analysé à différentes échelles de temps.

L'échelle accélérée $t \mapsto Nt$

C'est l'échelle qui s'avère être la plus intéressante pour étudier le comportement limite du système puisqu'elle capture à la fois son comportement à grande échelle et à long terme. Intuitivement, elle est

naturelle car, en accélérant le temps, elle permet de saisir le comportement stationnaire du processus lent $(V^N(t))$. Ainsi, si la valeur initiale du processus $(V_i^N(t))$ est d'ordre 1, c'est à dire que

$$\lim_{N\to+\infty} V_i^N(0) = v_i \in \mathbb{R},$$

on obtient que les processus $(V_i^N(t))$ et $(m_i^N(t))$ à l'échelle Nt sont négligeables devant N. Afin de préciser leur limite, définissons d'abord un outil adapté à l'étude de la convergence, dans un certain sens, des processus multi-échelles, la mesure d'occupation.

Définition 3.3.1 (*Mesure d'occupation*)

Soit g une fonction boorélenne positive à support compact dans $\mathbb{R}_+ \times \mathbb{N}^2$, la mesure d'occupation μ_N est définie par

$$\langle \mu_N, g\rangle = \frac{1}{N}\sum_{i=1}^{N} \int_{\mathbb{R}_+} g(u, m_i^N(Nu), V_i^N(Nu))du. \tag{3.3}$$

La présence inhabituelle du processus lent $(V_i^N(Nt))$ du nombre de clients de type 1 dans la mesure d'occupation nous permet de capturer la distribution conjointe limite du nombre d'emplacements disponibles et des clients de type 1. Cette distribution conjointe est le produit de deux géométriques dont les paramètres sont explicites. Le théorème suivant établit la comportement limite du système dans le cas sur-critique.

Théorème 3.3.2 (*Convergence des mesures d'occupation*)

Rappelons l'hypothèse $\lim_{N\to\infty} M/N = s$. Si $\beta c < \alpha$, la suite des mesures d'occupations (μ_N) sur l'espace d'état $\mathbb{R}_+ \times \mathbb{N}^2$ définies par l'equation (3.3) est tendue pour la convergence en distribution. Toute valeur d'adhérence μ_∞ satisfait

$$\langle \mu_\infty, g\rangle = \int_{\mathbb{R}_+} \mathbb{E}[g(u, X_1, X_2)]du \tag{3.4}$$

où g est une fonction à support fini dans $\mathbb{R}^+ \times \mathbb{N}^2$, X_1 et X_2 désignent des variables aléatoires indépendantes suivant des lois géométriques sur $\mathbb{N}$, de paramètres respectifs $\beta c/\alpha$ et ρ donné par

$$\rho = \frac{A + s + 1 - \sqrt{(A+s+1)^2 - 4sA}}{2A} \tag{3.5}$$

avec $A = \lambda\alpha/(\mu\beta c)$.

La preuve de ce résultat suit [Kur92] et se base sur plusieurs lemmes intermédiaires, notamment relatifs à la tension de la suite des mesures d'occupations (μ_N) et à la caractérisation donnée par (3.4) de toute valeur d'adhérence. Une étape clé de la preuve se fait en "jouant" avec l'équation d'évolution fonctionnelle (3.2), à l'échelle Nt, du processus de mesure empirique stochastique en normalisant une fois par N et une autre par N^2 avant de passer à la limite en faisant tendre N vers l'infini. Je la présente ici formellement. Considérons une fonction test g qui ne dépend pas de la première composante, c'est-à-dire que $g(u, m, v) = f(m, v)$, on peut réécrire, avec un abus de notations, la mesure d'occupation sous la forme suivante

$$\langle \mu_N, f\rangle = \frac{1}{N}\sum_{i=1}^{N} \int_0^t f(m_i(Nu), V_i(Nu))du = \int_0^t \Lambda^N(f)(Nu)du.$$

Considérons le cas où $f(m,v) = h(m)k(v)$, en normalisant par N^2, l'équation (3.2) devient, quand N tend vers l'infini

$$0 = \int_0^t \int_{\mathbb{N}^2} k(v) \left(\beta C \Delta_1(h)(m) + \alpha \Delta_{-1}(h)(m) \mathbf{1}_{m>0}\right) \pi_u(dm, dv) du$$

pour $t \geqslant 0$, avec des notations similaires à l'équation (3.1), $\Delta_i(h)(x) = h(x+i) - h(x)$, pour x et i deux entiers naturels. Ainsi, pour tout $t \geqslant 0$, on a

$$\int_{\mathbb{N}^2} k(v) \left(\beta c \Delta_1(h)(m) + \alpha \Delta_{-1}(h)(m) \mathbf{1}_{\{m>0\}}\right) \pi_t(dm, dv) = 0$$

qui peut être réécrit, pour tout $t \geqslant 0$,

$$\int_{\mathbb{N}^2} k(v) \Omega(h)(m) \pi_t(dm, dv) = 0$$

où, pour $h : \mathbb{N} \to \mathbb{R}^+$,

$$\Omega(h)(m) = \beta c \Delta_1(h)(m) + \alpha \Delta_{-1}(h)(m) \mathbf{1}_{\{m>0\}}.$$

Notons par $\pi_t(dm|v)$ la loi conditionnelle de m sachant v, et par $\pi_{2,t}(dv)$ la loi marginale de v. Alors

$$\int_{\mathbb{N}} \Omega(h)(m) \pi_t(dm|v) = 0$$

$\pi_{2,t}(dv)$ presque sûrement pour toute fonction h à support fini dans $\mathbb{N}$. Il suffit alors de voir que Ω est le générateur infinitésimal d'une file $M/M/1$ avec taux d'arrivée βc et taux de service α avec $\beta c < \alpha$. Par conséquence, $\pi_t(dm|v)$ est une loi géométrique de paramètre $\beta c/\alpha$ pour tout $v \geqslant 0$. Autrement dit

$$\int_{\mathbb{N}} \mathbf{1}_{\{m>0\}} \pi_t(dm|dv) = 1 - \pi_{1,t}(0) = \frac{\beta c}{\alpha}. \tag{3.6}$$

D'une manière similaire, et avec une fonction test g telle que $g(u,m,v) = k(v)$, l'équation (3.2), normalisée par N, devient, quand $N \to \infty$, pour tout $t \geqslant 0$,

$$0 = \int_{\mathbb{N}} \left(\lambda \Delta_{-1}(k)(v) \mathbf{1}_{v>0} + \mu \left(s - \int_{\mathbb{N}} v \ \pi_{2,u}(dv)\right) \Delta_1(k)(v) \frac{\beta C}{\alpha}\right) \pi_{2,t}(dv). \tag{3.7}$$

Dans le membre de droite de l'équation (3.7) apparaît le générateur infinitésimal d'un processus de vie et de mort avec comme taux de mortalité λ et taux de vie

$$\frac{\mu \beta c}{\alpha} \left(s - \int_{\mathbb{N}} v \ \pi_{2,t}(dv)\right).$$

La loi marginale $\pi_{2,t}(v)$ est alors une loi géométrique de paramètre

$$\rho = \frac{\mu}{\lambda} \frac{\beta c}{\alpha} \left(s - \int_{\mathbb{N}} v \ \pi_{2,t}(dv)\right).$$

Connaissant l'expression explicite de $\int_{\mathbb{N}} v \ \pi_{2,t}(dv)$ en fonction de ρ, l'équation précédente peut être réécrite comme une équation de point fixe en ρ dont la résolution est explicite

$$\rho = \frac{\mu}{\lambda} \frac{\beta C}{\alpha} \left(s - \frac{\rho}{1-\rho}\right).$$

L'échelle normale $t \mapsto t$

Dans cette échelle, lorsque la taille du système devient importante, le nombre de clients de type 1 sur chaque site se comporte comme un processus de Markov inhomogène appelé processus de McKean–Vlasov. La proposition 3 de [FMRT25] donne l'existence et l'unicité d'un tel processus, solution d'un système d'EDS. La preuve utilise un argument de point fixe tout à fait standard. Ce processus de Markov inhomogène a une belle interprétation en termes de files d'attente comme le nombre de clients $(L(t))$ d'une file $M/M/1$ avec un taux d'arrivée $(\beta c/\alpha)\mu\,(s - \mathbb{E}(L(t)))$ et taux de service λ. Dans ce régime sur-critique, l'impact de l'environnement est visible sur le taux d'arrivée modéré par la probabilité d'acceptation $\beta c/\alpha$. Le théorème suivant donne la convergence en distribution de $\left(\sum_{i=1}^{N} V_i^N(t)/N\right)$ à partir d'un certain $t_a \geqslant 0$ vers le processus $(L(t))$.

Théorème 3.3.3 (*Convergence champ-moyen*)

Supposons $\beta c < \alpha$, que pour tout $1 \leqslant i \leqslant N$, $\lim_{N\to\infty} V_i^N(0)/N = 0$ et que $\sup_i m_i^N(0) \leqslant aN$. Alors, il existe $t_a \geqslant 0$ fini tel que, partant de t_a, le processus $(\|V^N(t)\|/N) := \left(\sum_{i=1}^{N} V_i^N(t)/N\right)$ converge en distribution vers un processus $(L(t))$ identifié au nombre de clients dans une file $M/M/1$ avec taux d'arrivée $(\beta c/\alpha)\mu\,(s - \mathbb{E}(L(t)))$ et taux de service λ.

3.4 Retour vers le *free-floating*

La motivation de cette étude étant les systèmes d'autopartage sans stations physiques, appelés *free-floating*, il est alors naturel de vouloir tenir compte de la réservation de la voiture par l'usager avant d'effectuer son trajet. Techniquement, ceci revient à modifier les dynamiques du modèle en ajoutant un temps de réservation suivant une loi exponentielle de paramètre ν. Le processus décrivant l'état des N pseudo-stations devient alors

$$((m_i^N(t), V_i^N(t), R_i^N(t)),\ 1 \leqslant i \leqslant N)$$

où $m_i^N(t)$ est le nombre de places de parking disponibles, $V_i^N(t)$ le nombre des voitures partagées disponibles et $R_i^N(t)$ le nombre des voitures partagées réservées à la pseudo-station i pour $i = 1, ..., N$ à l'instant t, le nombre des voitures privées $X_i^N(t)$ dans une pseudo-station i donnée étant alors déduit grâce à la conservation de masse

$$m_i^N(t) + V_i^N(t) + R_i^N(t) + X_i^N(t) = cN.$$

Comportement limite du système

Pour ce modèle étendu, la même transition de phase est observée lorsque $\beta c = \alpha$ indépendamment des paramètres des voitures partagées. Ainsi, il existe deux régimes.

- Le régime de forte charge : Lorsque $c < \alpha/\beta$, le taux d'arrivée αN des voitures privées dépasse le taux d'achèvement du stationnement βcN. Par conséquent, il y a une probabilité positive que les voitures partagées ne puissent pas être garées dans la zone de destination et doivent être garées dans une autre zone. Cette probabilité, appelée probabilité de blocage, est un indicateur de l'inconfort ressenti par l'usager en cas de saturation.
- Le régime de faible charge : Lorsque $c > \alpha/\beta$, les voitures particulières occupent grosso modo une proportion $(\alpha/\beta)N < cN$ des places de parking avec une fluctuation d'ordre $\sqrt{N}$. Cela permet aux voitures partagées d'être garées à tout moment puisque leur nombre est d'ordre 1.

Quand le système est en surcharge, l'analyse multi-échelle de ce modèle étendu nous fournit deux résultats similaires à ceux obtenus plus haut sans la réservation.

Théorème 3.4.1 (*Comportement du système à l'échelle accélérée*)

Si $\beta c < \alpha$, quand N tend vers $+\infty$ avec $\lim_{N\to\infty} M/N = s$, la suite des mesures d'occupation (μ_N) sur l'espace d'état $\mathbb{R}_+ \times \mathbb{N}^3$ definies par l'équation (3.3) *est tendue pour la convergence en distribution. Toute valeur d'adhérence μ_∞ satisfait*

$$\langle \mu_\infty, g \rangle = \int_{\mathbb{R}_+} \mathbb{E}[g(u, X_1, X_2, X_3)]du \tag{3.8}$$

où g est une fonction à support fini dans $\mathbb{R}^+ \times \mathbb{N}^3$, X_1, X_2 et X_3 sont des variables aléatoires indépendantes avec X_1, X_2 suivant des lois géométriques sur $\mathbb{N}$ de paramètres respectifs $\beta c/\alpha$ et ρ alors que X_3 suit une loi de Poisson de paramètre $\rho\lambda/\nu$, où ρ est donné par l'équation (3.5) *où*

$$A = \lambda\left(\frac{\alpha}{\beta C} + \frac{1}{\nu}\right). \tag{3.9}$$

Le théorème 3.4.1 dit que, dans un certain sens, quand le système est en surcharge, le nombre de zones sans places de parking disponibles se comporte comme une distribution géométrique de paramètre $\beta c/\alpha$, celui des voitures partagées disponibles comme une distribution géométrique de paramètre ρ et celui des voitures partagées réservées comme une distribution de Poisson de paramètre $\rho\lambda/\nu$ et ce pour N grand. Ainsi, l'usager d'une voiture partagée est confronté, avec une probabilité non nulle $1 - \beta c/\alpha > 0$, à la possibilité de ne pas trouver de place de parking disponible et donc de devoir chercher une place dans une autre zone. Contrairement au cas d'un système d'autopartage basé sur des stations physiques, cette probabilité de blocage ne dépend que de l'environnement, c'est-à-dire des paramètres des voitures particulières α et β et du paramètre de la zone urbaine c. L'opérateur ne peut donc pas agir sur cette probabilité.
Pour l'échelle normale $t \mapsto t$, le résultat suivant est une extension immédiate du Théorème 3.3.3.

Théorème 3.4.2 (*Comportement du système à l'échelle normale*)

Si $\beta c < \alpha$ et $\sup_i m_i^N(0) \leqslant aN$, alors il existe $t_a > 0$ tel que, pour i fixé, $\left((V_i^N(t), R_i^N(t))/N, t > t_a\right)$ converge en distribution quand N tend vers $+\infty$ vers le processus inhomogène $(V(t), R(t))$ du nombre de clients dans un tandem de deux files d'attente (voir Figure 3.1) avec

- *la première file a un serveur avec taux d'arrivée $(\beta c/\alpha)\mu\left(s - \mathbb{E}\left(V(t) + R(t)\right)\right)$ à l'instant t et taux de service λ,*
- *et la deuxième a une infinité de serveurs avec taux de service ν.*

Il est alors immédiat d'obtenir que ce processus de Markov inhomogène a une unique mesure invariante caractérisant ainsi son comportement à long terme. C'est le but du résultat suivant.

Proposition 3.4.3 (*Mesure stationnaire de $(V(t), R(t))$*)

Le processus $(V(t), R(t))$ admet une unique mesure invariante de la forme

$$Geo(\rho) \otimes Poisson\left(\frac{\lambda\rho}{\nu}\right) \tag{3.10}$$

où ρ est donné par l'équation (3.5) *avec*

$$A = \lambda\left(\frac{\alpha}{\beta C} + \frac{1}{\nu}\right).$$

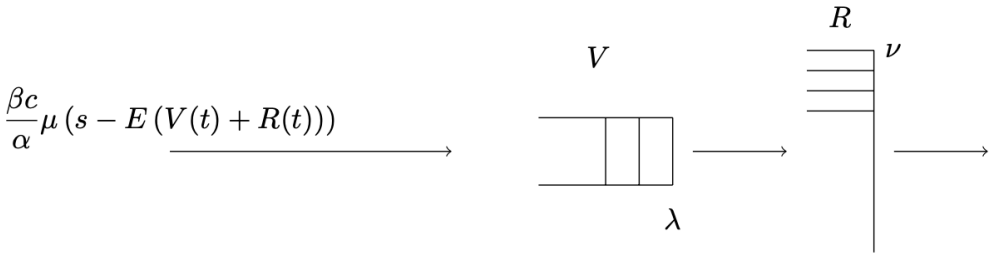

Figure 3.1 – Dynamique du processus limite $(V(t), R(t))$ en tant que tandem de deux files d'attente. La première file est une file $M/M/1$, la seconde une file $M/M/\infty$.

Cette mesure invariante est celle de (X_2, X_3) donné par le Théorème 3.4.1.

Le problème du dimensionnement

Rappelons que la proportion limite de pseudo-stations sans place de parking disponible est donnée par $1 - \beta c/\alpha > 0$, et donc ne dépend pas des paramètres des voitures partagées. L'opérateur ne peut donc pas agir sur cette proportion. Par contre, il peut agir sur la proportion de zones sans voitures partagées disponibles. Ceci est développé dans le résultat suivant.

Corollaire 3.4.4 (*Dimensionnement*)

Quand le système est en surcharge, la proportion limite $P_0 = \mathbb{P}(X_2 = 0)$ de zones sans voitures partagées disponibles, donnée par

$$P_0 = 1 - \rho \tag{3.11}$$

est une fonction croissante de s, avec ρ donné par l'équation (3.5) *et A défini par* (3.9).

Le corollaire 3.4.4 signifie que plus l'opérateur augmente le nombre de voitures partagées par pseudo-station, plus le système est performant, puisque la proportion de pseudo-stations sans voitures partagées disponibles diminue. Toutefois, ce résultat est limité par l'hypothèse selon laquelle le nombre de voitures partagées par pseudo-station est négligeable (de l'ordre de 1) par rapport au nombre de voitures particulières par pseudo-station (de l'ordre de N). Cette hypothèse est réaliste, étant donné l'ordre de grandeur du nombre de voitures privées par rapport à la taille de la flotte d'un système d'autopartage dans une ville.

Il est significatif que notre étude théorique puisse prouver que, même si l'opérateur double la taille de la flotte, par exemple, la probabilité de saturer l'espace public ne varie pas, puisqu'elle ne dépend pas des paramètres des voitures partagées. De plus, en augmentant la taille de la flotte, l'opérateur réduit la probabilité que l'usager ne trouve pas de voiture partagée disponible.

Il est aussi à noter que cette propriété de monotonie donnée par le corollaire 3.4.4 diffère de la forme en U de la courbe de performance (voir la Figure 2.5 par exemple), c'est à dire la courbe paramétrique $\rho \mapsto (s(\rho), \pi(\rho))$ exprimant la proportion de stations problématiques en fonction du ratio de dimensionnement s. L'effet de la capacité finie des pseudo-stations n'est pas visible pour les voitures partagées en *free-floating*.

3.5 Pour aller plus loin : routage pour une politique incitative

Nous venons de voir que l'opérateur n'a pas la main sur la probabilité pour qu'un usager trouve une place de parking disponible, il ne peut agir que sur la probabilité de trouver une voiture partagée disponible. Et pourtant, une constatation est faite par l'opérateur : même en augmentant la taille de la flotte, l'usager se plaint souvent de ne pas trouver de voiture partagée disponible. Pour répondre à cette problématique, une piste est la mise en place par l'opérateur d'une politique incitative afin que l'usager dépose à la fin de son trajet la voiture partagée dans une pseudo-station où il y a peu de voitures partagées disponibles. La nouveauté vient alors du mécanisme de routage qui n'est plus uniforme entre les pseudo-stations. Ainsi, la politique de routage dépend de l'état et, plus précisément, elle donne la priorité aux pseudo-stations où il y a peu ou pas de voitures partagées plutôt qu'aux pseudo-stations où il y en a davantage. Deux choix pour le mécanisme de routage sont analysés. Ils conduisent à deux limites champ-moyen différentes pour le système.

Le modèle

Reprenons le modèle *free-floating* présenté en section 3.2 où seul le choix de la destination est modifié comme suit : à la fin du trajet, le client de type 1 choisit le prochain nœud à visiter, disons le nœud j, non pas uniformément, mais selon la probabilité $P_j(v) = p(v_j)/\sum_{l=1}^{N} p(v_l)$. S'il ne trouve pas de place disponible au nœud j, il commence un nouveau trajet dont la durée a une distribution exponentielle de même paramètre μ et choisit ensuite une destination k parmi les N nœuds selon la probabilité $P_k(v)$ avec le même procédé. Cette opération est répétée jusqu'à ce que le client de type 1 trouve un nœud de destination avec une place disponible. Les poids $(p(v))$ sont pour le moment assez quelconques. En utilisant une approche champ-moyen similaire à ce qui précède, on obtient la loi marginale π_2 du nombre de voitures partagées disponibles dans une pseudo-station donnée. En effet, pour $k \in \mathbb{N}^+$, on a

$$\pi_2(k) = \pi_2(0) \left(\frac{\mu}{\lambda} \frac{\beta c}{\alpha} \frac{s - \int_{\mathbb{N}} v \ \pi_2(dv)}{\int_{\mathbb{N}} p(v) \ \pi_2(dv)} \right)^k \prod_{i=1}^{k} p(i-1).$$

Reste alors à tester quelques politiques de routage.

Deux mécanismes de routage

Ci-dessous deux choix de poids où la loi marginale π_2 du nombre de voitures partagées disponibles dans une pseudo-station donnée est explicite donc simple à manipuler. Ce sont des modèles "jouets" permettant à l'opérateur d'ajuster l'incitation qu'il veut mettre en place pour avantager, au moment du choix de la destination, les pseudo-stations en pénurie de voitures partagées, et donc favoriser une meilleure dispersion de la flotte sur l'ensemble de la zone de service.

Première politique

Considérons comme poids $p(k) = 1/(k+1)$ où k désigne le nombre de voitures partagées disponibles dans la pseudo-station en question. Le nombre de voitures partagées, à l'équilibre, a alors la distribution suivante. Pour $k \in \mathbb{N}$,

$$\pi_2(k) = \pi_2(0) \left(\frac{\mu}{\lambda} \frac{\beta c}{\alpha} \frac{s - \int_{\mathbb{N}} v \ \pi_2(dv)}{\int_{\mathbb{N}} \frac{1}{v+1} \ \pi_2(dv)} \right)^k \frac{1}{k!}$$

où on reconnaît une distribution de Poisson de paramètre

$$\rho := \frac{\mu}{\lambda}\frac{\beta c}{\alpha}\frac{s - \int_{\mathbb{N}} v\ \pi_2(dv)}{\int_{\mathbb{N}} \frac{1}{v+1}\ \pi_2(dv)}. \tag{3.12}$$

Ainsi, en réécrivant le terme de droite de (3.12) en terme de ρ, on obtient l'équation de point fixe qui suit pour $\rho > 0$ avec $A = \mu\beta c/(\lambda\alpha)$,

$$1 - e^{-\rho} + A\rho = As.$$

Seconde politique

Considérons un autre mécanisme de routage où le poids $p_r(k)$ dépend d'un certain paramètre $r \in \mathbb{N}^*$. Pour tout $k \in \mathbb{N}$,

$$p(k) = \frac{k+r}{k+1} = 1 + \frac{r-1}{k+1}$$

Le terme $(r-1)/(k+1)$ reflète un certain découragement des usagers à rejoindre une pseudo-station donnée, inversement proportionnel au nombre de voitures partagées disponibles dans cette pseudo-station. Le paramètre r exprime le niveau d'incitation que l'opérateur propose aux usagers. Lorsque $r = 1$, il n'y a pas d'incitation, le choix de destination est uniforme entre les nœuds et on retrouve le modèle initial. Le niveau d'incitation augmente avec r. À l'équilibre, la loi marginale π_2 du nombre de voitures partagées à une pseudo-station donnée est telle que, pour $k \in \mathbb{N}$,

$$\pi_2(k) = \pi_2(0)\left(\frac{\mu}{\lambda}\frac{\beta c}{\alpha}\frac{s - \int_{\mathbb{N}} v\ \pi_2(dv)}{\int_{\mathbb{N}} \frac{1}{v+1}\ \pi_2(dv)}\right)^k \frac{(k+r-1)!}{k!(r-1)!}.$$

On reconnaît alors une binomiale négative de paramètres r et ρ avec

$$\rho := \frac{\mu}{\lambda}\frac{\beta c}{\alpha}\frac{s - \int_{\mathbb{N}} v\ \pi_2(dv)}{\int_{\mathbb{N}} \frac{1}{v+1}\ \pi_2(dv)}.$$

En réécrivant le membre de droite de (3.5) en termes de ρ, on obtient l'équation de point fixe qui suit. Pour $\rho \in (0, 1]$ avec $A = \mu\beta c/(\lambda\alpha)$, on a

$$(1-\rho)^{r+1} + \rho(1 - As - Ar) = 1 - As.$$

Performance

Pour l'opérateur, le seul indicateur de performance est la proportion de zones sans voitures partagées disponibles, appelée aussi *probabilité de défaillance*. Dans notre modélisation, elle est approchée par $\pi_2(0)$ la probabilité stationnaire d'absence de voitures partagées dans une pseudo-station donnée. Ainsi, en se donnant un seuil, disons ε, pour cette probabilité de défaillance, on peut calculer la taille minimale de la flotte s telle que $\pi_2(0) \leq \varepsilon$. Ce problème peut être résolu explicitement pour les deux mécanismes de routage ci-dessus, le paramètre ρ est obtenu numériquement comme solution d'équation de point fixe.

Chapitre 4

Retour vers les structures discrètes : l'arbre exponentiellement préférentiel

Collaborations avec Rafik Aguech, Hosam Mahmoud, Toshio Nakata et Zhou Yang

L'arbre récursif est une structure hiérarchique classique. Des dizaines d'articles de recherche ont été consacrés au seul modèle uniforme, et nombre d'entre eux sont recensés dans [SM95]. Aujourd'hui, l'arbre récursif uniforme est un classique des structures aléatoires [Drm09, FK16, HM18]. Il est utilisé comme modèle dans de nombreuses applications telles que les systèmes pyramidaux [GB84] et la philologie [NH82]. D'autres applications ont suscité l'intérêt pour les modèles non uniformes dans lesquels l'insertion de nouveaux nœuds se fait de manière préférentielle en fonction d'un certain critère. Le premier de ces modèles préférentiels est un schéma de probabilités dans lequel les nœuds de degrés supérieurs sont favorisés [Szy87]. D'autres idées préférentielles sont basées sur l'âge [HM18, LM20] par exemple ou sur des poids de type puissance associés aux nœuds [LM22].
Dans [AMMY25], nous proposons un nouveau modèle préférentiel paramétré par une base (ou radix), un réel positif noté a. Dans ce modèle exponentiel, l'affinité d'un nœud est la base élevée à son étiquette. Ainsi, si la base est inférieure à 1, les nœuds qui apparaissent en premier (plus anciens dans l'arbre) ont un pouvoir d'attraction plus important que les nœuds plus récents. On observe ce phénomène dans la croissance des réseaux, où les nœuds les plus anciens ont plus de chances de croître que les plus jeunes.

4.1 Le modèle

On considère un arbre récursif où un nœud i à l'instant $n-1$ recrute avec une probabilité proportionnelle à a^i, avec a une constante réelle positive. Si on note par $A_{n,i}$ l'évènement que le nœud i recrute le nœud portant l'étiquette $n+1$ quand l'arbre a déjà n nœuds (l'arbre est d'age $n-1$), on obtient que

$$\mathbb{P}(A_{n,i}) = \frac{a^i}{\sum_{j=1}^{n} a^j} = \begin{cases} \dfrac{1}{n} & \text{si } \ a=1, \\ \dfrac{(a-1)\,a^{i-1}}{a^n-1} & \text{sinon.} \end{cases}$$

Un tel arbre sera appelé *arbre exponentiellement préférentiel avec base* a. Quand $a=1$, on retrouve le cas particulier de l'arbre récursif uniforme largement étudié [BFS92, Drm09, HM18, FK16, SM95]. La Figure 4.1 illustre les six arbres exponentiellement préférentiels de taille 4 avec comme base $a=1/2$. Les valeurs au-dessus des arbres sont leurs probabilités. Notez la forte probabilité attribuée

à l'arbre le plus "touffu" à l'extrême droite. Dans le modèle uniforme, cet arbre n'a qu'une probabilité de $1/6$.

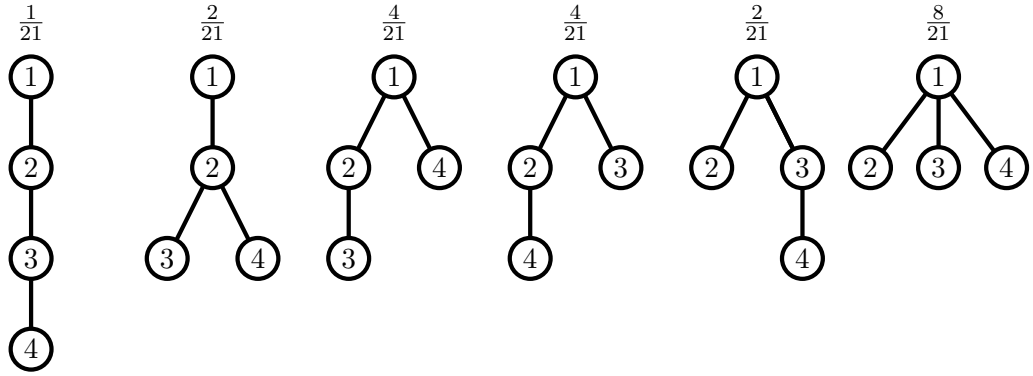

Figure 4.1 – Arbres exponentiellement préférentiels de taille 4 et base $a = 1/2$ avec leurs probabilités respectives.

Pour comprendre comment on obtient ces probabilités, je détaille ici le calcul d'un des arbres obtenus, le deuxième à partir de la gauche (dont la probabilité est $2/21$). Au départ, nous avons un nœud racine étiqueté 1. Avec une probabilité de 1, ce nœud racine recrute le nœud étiqueté 2. Ainsi, l'arbre ci-dessous apparaît avec probabilité 1. Les nœuds 1 et 2 sont maintenant en compétition pour attirer

le nœud 3, avec les probabilités suivantes $\frac{(1/2)^1}{(1/2)^1+(1/2)^2} = 2/3$ et $\frac{(1/2)^2}{(1/2)^1+(1/2)^2} = 1/3$. L'arbre

émerge avec probabilité $1 \times 1/3$. Les nœuds étiquetés 1, 2 et 3 sont maintenant en compétition pour attirer le nœud étiqueté 4, avec des probabilités respectives, $\frac{(1/2)^1}{(1/2)^1+(1/2)^2+(1/2)^3} = 4/7$, $\frac{(1/2)^2}{(1/2)^1+(1/2)^2+(1/2)^3} = 2/7$, et $\frac{(1/2)^3}{(1/2)^1+(1/2)^2+(1/2)^3} = 1/7$. Par conséquent, si le nœud étiqueté 2 est celui qui recrute, nous obtenons le deuxième arbre partant de gauche dans la Figure 4.1 avec la probabilité $1 \times 1/3 \times 2/7 = 2/21$. Cet exemple illustre la nature dynamique de la probabilité d'attraction au nœud i. Alors que a^i est un nombre fixe, l'affinité du nœud i évolue avec l'arbre.

4.2 Différents régimes

Intuitivement, on s'attend à une dichotomie relative à a, avec différents comportement de l'arbre, selon que $a < 1$ ou $a > 1$ et une transition de phase à $a = 1$. En effet, si $0 < a < 1$, alors les nœuds précoces sont les plus attractifs. Le cas $a > 1$ correspond au cas où les nœuds tardifs le sont.

Afin de visualiser ces différents régimes, on génère aléatoirement des arbres exponentiellement préférentiels de taille 100 avec différentes bases $a = 1/2$, 1 ou 2. On obtient la Figure 4.2 où la racine de chaque arbre est représentée par une étoile rouge. On appellera le régime $0 < a < 1$ *sous-critique*, le régime $a = 1$ *uniforme* et celui $a > 1$ *sur-critique*.

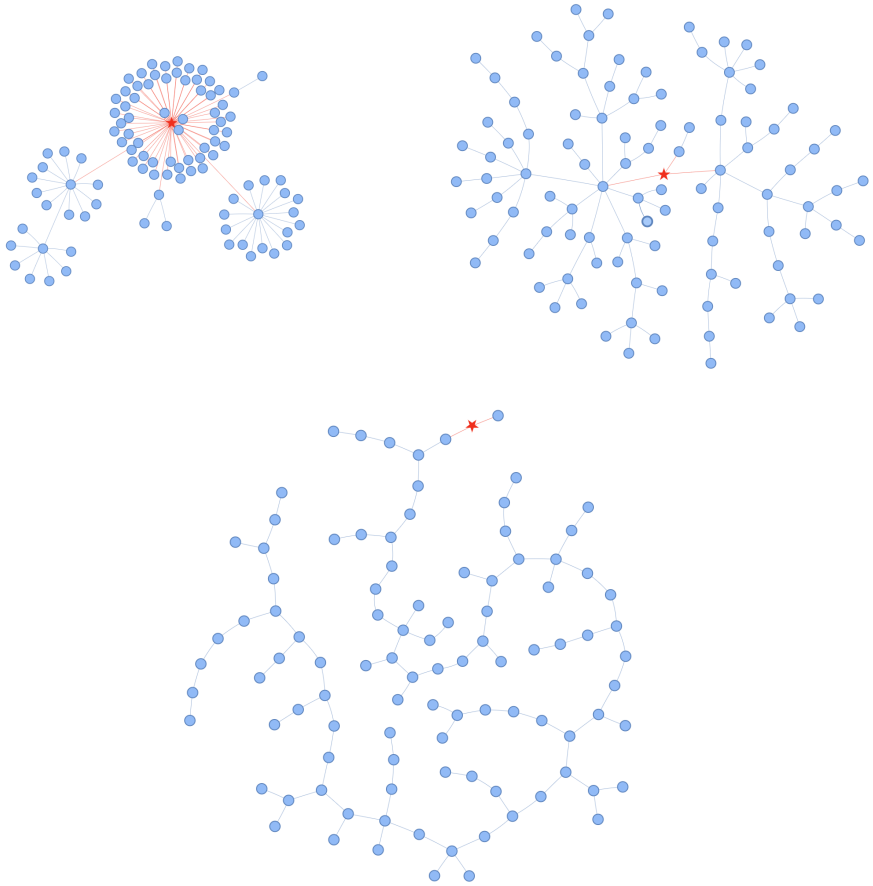

Figure 4.2 – Arbres de taille 100 générés aléatoirement : sous-critique (haut à gauche) avec $a = 1/2$, uniforme (haut à droite) avec $a = 1$, sur-critique (en bas) avec $a = 2$.

On remarque tout de suite que dans l'arbre sous-critique (celui en haut à gauche de la Figure 4.2), les nœuds sont regroupés près de la racine, ce qui en fait une structure d'arbuste. Dans l'arbre uniforme (en haut à droite de la Figure 4.2), les nœuds sont dispersés, tandis que dans l'arbre sur-critique (en bas de la Figure 4.2), de nombreux nœuds entraînent l'arbre vers des altitudes plus élevées, ce qui en fait un arbre grand et maigre, avec des branches courtes jaillissant d'un tronc principal mince.

Pour étudier ces différents régimes, on considère le degré sortant $\Delta_{n,i}$ du nœud i dans un arbre de taille n. Le degré sortant d'un nœud i augmente quand il recrute un autre nœud et on obtient la

représentation suivante

$$\Delta_{n,i} = \sum_{k=i}^{n-1} \mathbf{1}_{A_{k,i}}. \tag{4.1}$$

Le cas uniforme $a = 1$ étant bien étudié, je ne présente ici que les deux autres régimes, sous-critique ($a < 1$) et sur-critique ($a > 1$). À partir de l'équation (4.1), on établit l'expression exacte de la moyenne

$$\mathbb{E}[\Delta_{n,i}] = \sum_{k=i}^{n-1} \mathbb{P}(\mathbf{1}_{A_{k,i}} = 1) = \sum_{k=i}^{n-1} \frac{(a-1)\, a^{i-1}}{a^k - 1}, \tag{4.2}$$

pour $a \neq 1$. Quand n tend vers l'infini, on obtient les équivalents asymptomatiques suivants.

Proposition 4.2.1

Soit $\Delta_{n,i}$ le degré sortant d'un nœud i dans un arbre exponentiellement préférentiel de taille n et de base $0 < a < 1$. Pour n assez grand, on a

$$\mathbb{E}[\Delta_{n,i}] = (1-a)\, a^{i-1}\, (n-i) + O(a^i).$$

Si $a > 1$, alors deux cas sont à distinguer selon la relation entre i et n et on obtient pour n assez grand deux phases :

$$\mathbb{E}[\Delta_{n,i}] = \begin{cases} (a-1)\, a^{i-1} c_i^{(1)}(a) + O(\frac{1}{a^{n-i}}) & \text{si } i \text{ est fixé,} \\ 1 - \frac{1}{a^{n-i}} + O(\frac{1}{a^i}) & \text{si } n \geq i \to \infty \end{cases}$$

où

$$c_i^{(1)}(a) = \sum_{k=i}^{\infty} \frac{1}{a^k - 1}.$$

Pour obtenir la variance de $\Delta_{n,i}$, partons toujours de (4.1) où les indicatrices $\mathbf{1}_{A_{n,i}}$, pour $n \geq 1$ fixé, sont indépendantes. On obtient alors que

$$\mathrm{Var}[\Delta_{n,i}] = \mathrm{Var}\Big[\sum_{k=i}^{n-1} \mathbf{1}_{A_{k,i}}\Big] = \sum_{k=i}^{n-1} \mathrm{Var}[\mathbf{1}_{A_{k,i}}].$$

Des arguments combinatoires nous permettent d'établir un comportement asymptotique similaire à celui de la moyenne.

Proposition 4.2.2

Soit $\Delta_{n,i}$ le degré sortant du nœud i dans un arbre exponentiellement préférentiel de taille n et base $0 < a < 1$. On a

$$\mathrm{Var}[\Delta_{n,i}] = (1-a)\, a^{i-1}\big(1 - (1-a)\, a^{i-1}\big)(n-i) + O(a^i).$$

Si $a > 1$, alors deux cas sont à distinguer selon la relation entre i et n et on obtient pour n assez grand deux phases :

$$\mathrm{Var}[\Delta_{n,i}] = \begin{cases} (a-1)\, a^{i-1} c_i^{(1)}(a) - (a-1)^2\, a^{2i-2} c_i^{(2)}(a) + O(\frac{1}{a^{n-i}}), & \text{si } i \text{ est fixé,} \\ \frac{2}{a+1} - \frac{1}{a^{n-i}} + \frac{a-1}{(a+1)a^{2n-2i}} + O(\frac{1}{a^i}), & \text{si } n \geq i \to \infty \end{cases}$$

où

$$c_i^{(2)}(a) = \sum_{k=i}^{\infty} \frac{1}{(a^k - 1)^2}.$$

On peut alors en déduire un premier résultat relatif à la distribution de $\Delta_{n,i}$.

Corollaire 4.2.3

Dans le régime sous-critique ($0 < a < 1$), si $(n-i)a^i \to \infty$, quand n tend vers l'infini, alors on obtient que

$$\frac{\Delta_{n,i}}{(n-i)\,a^i} \xrightarrow{\mathbb{P}} \frac{1-a}{a}.$$

4.3 Différentes phases pour chaque régime

Quand on regarde de plus près, pour n large, la distribution du degré sortant $\Delta_{n,i}$ pour un nœud i, la relation entre i et n s'avère intrigante et fait apparaître différents comportements au sein d'un même régime : comportement gaussien, approximation poissonienne avec un phénomène d'oscillation ou convergence (en probabilités ou presque sûre). Le théorème de Barbour et Holst [BH84] permet d'obtenir des approximations de Poisson.

Théorème 4.3.1 (*Régime sous-critique*)

Soit $\Delta_{n,i}$ le degré sortant du nœud i dans un arbre exponentiellement préférentiel de taille n et base $0 < a < 1$. Alors, on observe trois phases :

(a) Une première phase, que nous appelons précoce *où $1 \leq i \leq \lfloor \log_{\frac{1}{a}} n \rfloor - g(n)$ quand $n \to \infty$ avec g une fonction positive à valeurs dans $\mathbb{N}$ allant vers l'infini, telle que $g(n) = o(\ln n)$. On obtient alors que*

$$\frac{\Delta_{n,i} - \frac{1-a}{a}\,a^i\,n}{\sqrt{a^i\,n}} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \frac{1-a}{a}\Big).$$

(b) Une seconde phase, appelée intermédiaire, *où $i = \lfloor \log_{\frac{1}{a}} n \rfloor + c$ pour $c \in \mathbb{Z}$, on a*

$$d_{TV}\left(\Delta_{n,i}, Poisson\Big(\frac{1-a}{a}\,a^{c-\{\log_{\frac{1}{a}} n\}}\Big)\right) \to 0,$$

où $\{x\}$ désigne la partie fractionnaire d'un réel x définie par $\{x\} = x - \lfloor x \rfloor$.

(c) La dernière phase, appelée tardive, *où $i = \lfloor \log_{\frac{1}{a}} n \rfloor + h(n)$ avec h une fonction positive, non constante, à valeurs dans $\mathbb{N}$ allant vers l'infini telle $h(n) = O(n - \lfloor \log_{\frac{1}{a}} n \rfloor)$. On obtient alors que*

$$\Delta_{n,i} \xrightarrow{\mathbb{P}} 0.$$

Le cas critique ($a = 1$) est similaire à celui sous-critique en terme de comportement de l'arbre pour n grand. En effet on obtient trois phases : comportement gaussien, approximation poissonnienne et limite nulle. Pour ce cas classique, je renvois vers [Kar93, MS93]. Bien qu'il présente aussi trois phases, le régime sur-critique se distingue en terme de comportement limite. Les calculs et arguments combinatoires sont plus compliqués à manipuler.

Théorème 4.3.2 (*Régime sur-critique*)

Pour le régime sur-critique ($a > 1$), trois phases apparaissent :

(a) Une phase précoce *où i est fixé, on obtient quand $n \to \infty$, que $\Delta_{n,i} \xrightarrow{p.s.} \Delta_i^*$, et pour $k \geq 1$, la variable aléatoire limite a pour distribution*

$$\mathbb{P}(\Delta_i^* = k) = (a-1)^k a^{(i-1)k} \lim_{n\to\infty} \Big(\Big(\prod_{\ell=i}^{n-1} \frac{1}{a^\ell - 1} \Big) \sum_{i \leq j_1 < j_2 < \cdots < j_k \leq n-1} \times \prod_{\substack{i \leq m \leq n-1 \\ m \notin \{j_1,\ldots,j_k\}}} \Big((a^m - 1) - (a-1)a^{i-1} \Big) \Big).$$

(b) Une phase intermédiaire *où $1 \leq i = n - b(n)$ avec b une fonction allant vers l'infini, telle que $n - b(n) \to +\infty$ aussi. On obtient alors que*

$$\mathbb{P}(\Delta_{n,n-b(n)} = k) = (a-1)^k a^{(n-b(n)-1)k} \Big(\prod_{\ell=i}^{n-1} \frac{1}{a^\ell - 1} \Big) \times \sum_{n-b(n) \leq j_1 < j_2 < \cdots < j_k \leq n-1} \prod_{\substack{i \leq m \leq n-1 \\ m \notin \{j_1,\ldots,j_k\}}} \big((a^m - 1) - (a-1)a^{n-b(n)-1} \big).$$

(c) Une phase tardive *où $1 \leq i = n - c$, pour $c \in \mathbb{N}$. On obtient alors que $\Delta_{n,n-c} \xrightarrow{p.s.} \Delta_c^\star$, où $\Delta_c^\star$ a pour distribution*

$$\mathbb{P}(\Delta_c^\star = k) = \frac{(a-1)^k}{a^{c(c+1)/2}} \sum_{1 \leq r_1 < r_2 < \cdots < r_k \leq c} \prod_{\substack{1 \leq s \leq c \\ s \notin \{r_1,\ldots,r_k\}}} (a^s - a + 1).$$

Rappelons que i est l'étiquette d'un nœud donné, il sera toujours considéré comme un entier naturel. Il faut aussi remarquer que les mots *précoce*, *intermédiaire* et *tardif* n'ont pas la même signification selon les différents régimes.

4.4 Quelques remarques et pistes futures

Deux exemples illustratifs

Ci-dessous deux exemples du régime sur-critique avec $a = 2$.

– La phase intermédiaire : L'expression exacte de la distribution du degré sortant d'un nœud i pour le régime sur-critique dans sa phase intermédiaire (Théorème 4.3.2 (b))), est peu pratique, mais elle peut être utilisée pour calculer la probabilité pour un petit k, par exemple, qu'un nœud intermédiaire soit une feuille ($k = 0$) ou apporter une information sur la structure asymptotique de l'arbre. Considérons l'exemple d'un nœud intermédiaire $i = n - \lfloor \ln n \rfloor$. Dans ce cas, $b(n)$ est donnée par $\lfloor \ln n \rfloor$. Pour $k = 0$, l'ensemble $\{j_1, \ldots, j_0\}$ est vide donc il ne reste qu'un seul produit sur la totalité de l'ensemble $\{n - h(n), \ldots, n-1\}$. On obtient alors que

$$\mathbb{P}(\Delta_{n,n-\lfloor \ln n \rfloor} = 0) = \frac{1}{\prod_{r=n-\lfloor \ln n \rfloor}^{n-1} (2^r - 1)} \prod_{m=n-\lfloor \ln n \rfloor}^{n-1} (2^m - 1 - 2^{n-\lfloor \ln n \rfloor - 1}).$$

Pour $n = 1000$, cette probabilité est d'environ 0.2933.

– Les nœuds les plus tardifs se taillent la part du lion : Considérons un arbre de taille n assez grand. Le degré sortant du nœud n est 0, car il n'a pas encore pu recruter d'autres nœuds. Le nœud $n-1$, dont l'étiquette est la deuxième plus élevée de l'arbre, ne peut recruter que le nœud n avec la probabilité $2^{n-1}(2-1)/(2^n-1) \to 1/2$. Ainsi, le degré sortant de l'avant-dernier nœud est asymptotiquement distribué comme une loi de Bernoulli de paramètre $1/2$.
Le nœud $n-2$ a deux possibilités de recrutement et donc une distribution asymptotique sur $\{0,1,2\}$ de moyenne $1/4$, et ainsi de suite. Selon le Théorème 4.3.2 (c), on a

$$\mathbb{P}(\Delta^\star_{n,n-2}=k) \to \mathbb{P}(\Delta^\star_2=k) = \frac{1}{2^3} \sum_{1\le r_1<r_2<\cdots<r_k\le 2} \prod_{\substack{1\le s\le 2\\ s\notin\{r_1,\ldots,r_k\}}} (2^s-1),$$

pour $k=0,1,2$. Pour $k=0$, l'ensemble $\{r_1,r_2,\ldots,r_0\}$ est vide, et on obtient

$$\mathbb{P}(\Delta^\star_2=0) = \frac{1}{8} \prod_{\substack{1\le s\le 2\\ s\notin\phi}} (2^s-1) = \frac{1\times 3}{8} = \frac{3}{8}.$$

De plus, on a

$$\mathbb{P}(\Delta^\star_2=1) = \frac{1}{8} \sum_{1\le k_1\le 2} \prod_{\substack{1\le s\le 2\\ s\notin\{k_1\}}} (2^s-1) = \frac{3+1}{8} = \frac{4}{8}.$$

Et donc $\mathbb{P}(\Delta^\star_2=2) = 1-\mathbb{P}(\Delta^\star_2=0)-\mathbb{P}(\Delta^\star_2=1) = 1/8$. Pour résumer, on obtient la distribution limite suivante

$$\Delta^\star_2 = \begin{cases} 0, & \text{avec probabilité } 3/8;\\ 1, & \text{avec probabilité } 4/8;\\ 2, & \text{avec probabilité } 1/8. \end{cases}$$

Remarques et travaux en cours

Remarque 4.4.1

Dans les Théorèmes 4.3.1 et 4.3.2, nous avons discuté des phases où la croissance de i vers n est systématique. Cependant, il n'y a pas de limite à la bizarrerie de la suite $i=i(n)$. Par exemple, $i(n)$ peut être une suite alternant entre deux (ou plusieurs valeurs), comme la suite $i(n)=5+(-1)^n$, dans laquelle le degré sortant du noeud i ne converge pas. Pire encore, $i(n)$ peut ne pas avoir de structure du tout.

Pour généraliser le modèle d'arbre exponentiellement préférentiel où l'attractivité d'un nœud i est proportionnelle à a^i, on se propose de considérer une suite de poids plus générale. Nakata et Mahmoud considèrent dans [NM24] un modèle assez général, dans lequel les poids suivent une suite arbitraire statique a_i de nombres réels positifs. Ils appellent (a_i) la suite d'affinité et considèrent a_i comme l'affinité du nœud i après son apparition dans l'arbre. Pour $n \geqslant 2$, notons par $A_{n,i}$ l'événement que le nœud i recrute le nœud n (lorsque la taille de l'arbre est $n-1$). Les affinités sont transformées en probabilités en normalisant par $s_n := \sum_{i=1}^n a_i$. Ainsi,

$$\mathbb{P}(A_{n,i}) = \frac{a_i}{s_{n-1}} \text{ pour } i=1,\ldots,n-1.$$

Le modèle analysé dans [NM24] représente une classe entière d'arbres récursifs avec des affinités générales. Il s'agit d'un modèle préférentiel *statique* où la suite (a_i) est fixée à l'instant i et reste la

même pendant toute la durée de vie de l'arbre contrairement à une version dynamique où l'affinité du nœud i est une fonction à la fois de i et de n. Il est intéressant de considérer des affinités de la forme $a_{n,i}$ dont un exemple est le modèle d'âge avec des affinités $a_{n,i} = n - i$. Ce modèle paraît assez naturel puisque les nœuds avec des étiquettes plus petites (donc des nœuds plus anciens dans l'arbre) ont des probabilités de recrutement plus élevées. Une telle affinité dynamique tient compte de l'*expérience*. Dans un travail en cours [AMMN25], on considère des affinités dynamiques de la forme

$$a_{n,i} = \alpha n + \beta i \text{ pour } i = 1, \ldots, n-1, \ \alpha \text{ et } \beta \text{ étant deux constantes.}$$

Dans ce travail, nous élargissons notre étude à d'autres propriétés de l'arbre, telles que le degré maximal, la profondeur des nœuds et la longueur totale de chemin.

Bibliographie

[AFM22] J. Ancel, C. Fricker, and H. Mohamed. Mean field analysis for bike and e-bike sharing systems. *ACM SIGMETRICS Performance Evaluation Review*, 49 :12–14, 01 2022.

[AMMN25] R. Aguech, H. Mahmoud, H. Mohamed, and T. Nakata. Recursive trees grown from dynamic building sequences. working paper, 2025.

[AMMY25] R. Aguech, H. Mahmoud, H. Mohamed, and Z. Yang. Exponentially preferential trees. acceptable for publication in the Network Science journal, 2025.

[ARS18] R. Aghajani, P. Robert, and W. Sun. A large scale analysis of unreliable stochastic networks. *The Annals of Applied Probability*, 28(2) :p. 851 – 887, 2018.

[BCMP75] F. Baskett, K.M. Chandy, R.R. Muntz, and F.G. Palacios. Open, closed, and mixed networks of queues with different classes of customers. *J. ACM*, 22(2) :248–260, 1975.

[BFS92] F. Bergeron, P. Flajolet, and B. Salvy. Varieties of increasing trees. *Lecture Notes in Computer Science*, 581 :24–48, 1992.

[BH84] A. Barbour and P. Hall. On the rate of poisson convergence. *Mathematical Proceedings of Cambridge Philosophical Society*, 95 :473–480, 1984.

[BMP10] C. Bordenave, D. McDonald, and A. Proutiere. A particle system in interaction with a rapidly varying environment : Mean field limits and applications. *Networks and Heterogeneous Media*, 5(1) :31–62, 2010.

[CBLW21] F. Cecchi, S.C. Borst, J.S. Leeuwaarden, and P.A. Whiting. Mean-field limits for large-scale random-access networks. *Stochastic Systems*, 11(13) :193–217, 2021.

[CBM+21] E. Castiel, S. Borst, L. Miclo, F. Simatos, and P. Whiting. Induced idleness leads to deterministic heavy traffic limits for queue-based random-access algorithms. *The Annals of Applied Probability*, 31(2) :941–971, 2021.

[DFM18] P. Dester, C. Fricker, and H. Mohamed. Stationary distribution analysis of a queueing model with local choice. In *Proceedings vol. AQ, 29th Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'18)*, 2018.

[Drm09] M. Drmota. Random trees : An interplay between combinatorics and probability. *Springer-Verlag Wien New York*, 2009.

[EF98] A. Economou and D. Fakinos. Product form stationary distributions for queueing networks with blocking and rerouting. *Queueing Systems Theory Appl*, 30 :251–260, 1998.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov Processes : Characterization and Convergence*. John Wiley & Sons Inc, 1986.

[FG16] C. Fricker and N. Gast. Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. *Euro journal on transportation and logistics*, 5(3) :261–291, 2016.

[FGM12] C. Fricker, N. Gast, and H. Mohamed. Mean field analysis for inhomogeneous bike sharing systems. In *Proceedings vol. AQ, 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12)*, 2012.

[FIM17] G. Fayolle, R. Iasnogorodski, and V. Malyshev. *Random Walks in the Quarter Plane*. Springer International Publishing, 2017.

[FJ07] C. Fricker and M.R. Jaïbi. On the fluid limit of the $M/G/\infty$ queue. *Queueing Systems*, 56 :255–265, 2007.

[FK16] A. Frieze and M. Karon'ski. *Introduction to Random Graphs*. 2nd Ed. Cambridge University Press, Cambridge, UK, 2016.

[FL96] G Fayolle and J-M Lasgouttes. Asymptotics and scalings for large product-form networks via the central limit theorem. *Markov Process Related Fields*, 2(2) :317–348, 1996.

[FM25] Christine Fricker and Hanene Mohamed. Mean-field analysis of stochastic networks with reservation. *Journal of Applied Probability*, pages 1–27, 01 2025.

[FMB20] C. Fricker, H. Mohamed, and C. Bourdais. A mean field analysis of a stochastic model for reservation in car-sharing systems. *ACM SIGMETRICS Performance Evaluation Review*, 48(2) :18–20, 2020.

[FMM95] G. Fayolle, V.A. Malyshev, and M.V. Men'shikov. Topics in the constructive theory of countable markov chains. *Cambridge university press*, 05 1995.

[FMR25] C. Fricker, H. Mohamed, and A. Rigonat. Stochastic averaging and mean-field for a large system with fast varying environment with applications to free-floating car-sharing. soumis, disponible sur HAL, 2025.

[FMRT25] C. Fricker, H. Mohamed, A. Rigonat, and M. Trépanier. A new stochastic model for carsharing suited to free-floating. *Transportation Research Procedia*, 82 :2395–2409, 2025.

[FRZ23] V. Fromion, P. Robert, and J. Zaherddine. Stochastic models of regulation of transcription in biological cells. *Journal of Mathematical Biology*, 87(5), 2023.

[FT17] C. Fricker and D. Tibi. Equivalence of ensembles for large vehicle-sharing models. *The Annals of Applied Probability*, 27(2) :883 – 916, 2017.

[Gas16] N. Gast. Construction of lyapunov functions via relative entropy with application to caching. In *The 18th Workshop on MAthematical performance Modeling and Analysis, Nice, France*, Jun 2016.

[GB84] J. Gastwirth and P. Bhattacharya. Two probability models of pyramid or chain letter schemes demonstrating that their promotional claims are unreliable. *Operations Research*, 32 :527–536, 1984.

[GB02] Godfrey A. Gregory and Powell Warren B. An adaptive dynamic programming algorithm for dynamic fleet management, ii : Multiperiod travel times. *Transportation Science*, 36(1) :40–54, 2002.

[GB10] N. Gast and G. Bruno. A mean field model of work stealing in large-scale systems. In *ACM SIGMETRICS Performance Evaluation Review*, volume 38, pages 13–24. ACM, 2010.

[GHBSW+17] K. Gardner, M. Harchol-Balter, A. Scheller-Wolf, M. Velednitsky, and S. Zbarsky. Redundancy-d : The power of d choices for redundancy. *Oper. Res.*, 65(4) :1078–1094, 2017.

[GHK90] R.J. Gibbens, P.J. Hunt, and F.P. Kelly. Bistability in communication networks. *Disorder in physical systems*, pages 113–128, 1990.

[GLM+10] A. Ganesh, S. Lilienthal, D Manjunath, A. Proutiere, and F. Simatos. Load balancing via random local search in closed and open systems. In *ACM SIGMETRICS Performance Evaluation Review*, volume 38, pages 287–298. ACM, 2010.

[GM93] C. Graham and S. Méléard. Propagation of chaos for a fully connected loss network with alternate routing. *Stochastic processes and their applications*, 44(1) :159–180, 1993.

[GX10] DK. George and CH. Xia. Asymptotic analysis of closed queueing networks and its implications to achievable service levels. *SIGMETRICS Performance Evaluation Review*, 38(2) :3–5, 2010.

[GX11] DK. George and CH. Xia. Fleet-sizing and service availability for a vehicle rental system via closed queueing networks. *European Journal of Operational Research*, 211(1) :198–207, 2011.

[HK94] P. Hunt and T. Kurtz. Large loss networks. *Stochastic Processes and their Applications*, 53(2) :363–378, 1994.

[HM18] M. Hofri and H. Mahmoud. Algorithmics of nonuniformity : Tools and paradigms. *CRC Press, Boca Raton, Florida*, 2018.

[Kac56] M. Kac. Foundations of kinetic theory. In *Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability*, volume 3, pages 171–197, 1956.

[Kar93] A. Karr. *Probability*. Springer-Verlag, New York, 1993.

[Kur92] T.G. Kurtz. Averaging for martingale problems and stochastic approximation. *Applied Stochastic Analysis*, pages 186–209, 1992.

[LM20] M. Lyon and H. Mahmoud. Trees grown under young-age preferential attachment. *Journal of Applied Probability*, 57 :911–927, 2020.

[LM22] M. Lyon and H. Mahmoud. Insertion depth in power-weight trees. *Information Processing Letters*, 176 :106227, 2022.

[McK66] H.P. McKean. A class of markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences of the United States of America*, 56(6) :1907–1911, 1966.

[MFM+22] B. Moreno, C. Fricker, H. Mohamed, A. Philippe, and M. Trépanier. Mean field analysis of an incentive algorithm for a closed stochastic network. In *Proceedings vol. AQ, 33rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'22)*, volume 13, pages 1–17, 2022.

[MFM+24] B. Moreno, C. Fricker, H. Mohamed, A. Philippe, and M. Trépanier. An incentive algorithm for a closed stochastic network : Data and mean-field analysis. *La Matematica*, 3 :651–676, 2024.

[Mit96] M. Mitzenmacher. *The Power of Two Choices in Randomized Load Balancing*. PhD thesis, Berkeley, 1996.

[MM79] V.A. Malyshev and M.V. Menshikov. Ergodicity, continuity, and analyticity of countable markov chains. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 39 :3–48, 1979.

[MS93] R. Mahmoud, H. Smythe and J. Szyman'ski. On the structure of plane-oriented recursive trees and their branches. *Random Structures & Algorithms*, 4 :151–176, 1993.

[MV96] Yakovlev A.V. Malyshev V.A. Condensation in large closed jackson networks. *The Annals of Applied Probability*, 6(1) :92–115, 1996.

[NH82] D. Najock and C. Heyde. On the number of terminal vertices in certain random trees with an application to stemma construction in philology. *Journal of Applied Probability*, 19 :675–680, 1982.

[NM24] T. Nakata and H. Mahmoud. Bernoulli convolution of the depth of nodes in recursive trees with general affinities. *Journal of Stochastic Analysis*, 5(2), 2024.

[Rob13] P. Robert. *Stochastic networks and queues*, volume 52. Springer, 2013.

[SM95] R. Smythe and H. Mahmoud. A survey of recursive trees. *Theory of Probability and Mathematical Statistics*, 51 :1–27, 1995.

[Szn89] A.S. Sznitman. Topics in propagation of chaos. *Ecole de Probabilités de Saint Flour, XIX-1989. Lecture Notes in Math.*, pages 165–251, 1989.

[Szy87] J. Szyman'ski. On a nonuniform random recursive tree, north-holland mathematics studies. *North-Holland Mathematics Studies*, 144 :297–306, 1987.

[VDK96] N. Vvedenskaya, R. Dobrushin, and F. Karpelevich. Queueing system with selection of the shortest of two queues : An asymptotic approach. *Problemy Peredachi Informatsii*, 32(1) :20–34, 1996.