

Projets MODAL MAP441  
Simulation numérique aléatoire (SNA)  
autour des évènements rares  
Promotion X2011

*Equipe pédagogique :*

F. Benaych-Georges, S. De Marco, S. Gaïffas, E. Gobet, C. Vergé

*Période : Mai 2013 - Juin 2013*

*(Version du 27 avril 2013 )*

## Table des matières

<b>1</b>	<b>Dimensionnement d'une mémoire tampon</b>	<b>3</b>
<b>2</b>	<b>Sûreté d'un réacteur nucléaire</b>	<b>7</b>
<b>3</b>	<b>Réaction chimique</b>	<b>10</b>
<b>4</b>	<b>Risque systémique</b>	<b>13</b>
<b>5</b>	<b>Messages twitter</b>	<b>17</b>
<b>6</b>	<b>Ruines en assurance</b>	<b>21</b>
<b>7</b>	<b>Barrages et risques d'inondation</b>	<b>25</b>
<b>8</b>	<b>Equation de la chaleur et résolution Monte-Carlo multi-niveaux (pour binôme mixte SNA-SNE)</b>	<b>29</b>
<b>9</b>	<b>Solutions approchées certifiées pour l'équation de Burgers visqueuse et incertitudes (pour binôme mixte SNA-SNE)</b>	<b>31</b>

**Objectifs.** Il s'agit de réaliser un projet de modélisation et simulation sur un des thèmes applicatifs décrits ci-dessous, dans le contexte des événements rares<sup>1</sup>.

Par la bibliographie proposée, on attend des étudiants qu'ils :

- se familiarisent avec le thème proposé,
- en comprennent les enjeux de sorte à se poser les *bonnes questions*,
- développent des outils stochastiques de simulation pour mieux appréhender les événements rares (quantile, distribution conditionnelle à l'évènement, statistique dans de tels événements, scénario rare typique...). Ils auront en particulier à adapter à leur contexte les méthodes abordés en cours ou en TP.

Cela implique de l'autonomie et des initiatives.

Il est demandé de tester différentes approches, de les *comparer quantitativement et qualitativement*, avec un regard critique.

**Encadrement :** une permanence de suivi de projets est assurée chaque vendredi après-midi en salles informatiques 30-31, de 13h30 à 15h30.

**Livrables :** un rapport écrit, avec les codes de simulation en annexe.

**Notation :** le projet est noté sur 20 pts. Il est recommandé de choisir *Scilab* comme langage de programmation. En cas de besoin en calculs particulièrement intensifs (par exemple usages inévitables de boucles), on pourra utiliser Java ou C/C++.

Une soutenance est prévue **mercredi 26 juin 2013** : la durée est de 45' (30' de présentation, 15' de questions).

**Barème de notation du projet :**

- (4pts) Capacités d'initiatives/autonomie face aux difficultés/problèmes soulevés par le projet.
- (4pts) Capacités à décomposer le problème en plusieurs étapes intermédiaires, avant de parvenir à une version aboutie (démarche d'expérimentation).
- (5pts) Capacités à mesurer les améliorations, les apports, comparaison numérique.
- (4pts) Rapport écrit.
- (3pts) Soutenance orale.

Les 3 premiers items sont principalement notés pendant les séances de TPs.

---

1. Des projets mixtes SNA-SNE sont également proposés (voir Projets 8 et 9), avec la contrainte que le binôme soit composé d'un élève inscrit en SNA, et l'autre en SNE.

# 1 Dimensionnement d'une mémoire tampon

## • Le modèle

Considérons l'émetteur d'un réseau de télécommunication fonctionnant de la façon suivante : des paquets (les messages) lui arrivent à des instants aléatoires, ils sont alors acheminés dans leur ordre d'arrivée. Le temps nécessaire à l'acheminement de chaque paquet est non négligeable. Il est donc possible qu'un paquet arrive sans pouvoir être traité tout de suite : avant d'être envoyés, les paquets sont alors stockés dans une mémoire, appelée *mémoire tampon* de l'émetteur. La mémoire tampon n'est pas de taille infinie, elle doit être dimensionnée de façon à ce que sa saturation soit très rare.

On modélise les instants d'arrivée des paquets avec un processus ponctuel de Poisson sur  $\mathbb{R}^+$  d'intensité  $\lambda > 0$  : les paquets arrivent à des instants  $T_1 < T_2 < T_3 < \dots$  tels que, en posant  $T_0 = 0$ , les intervalles de temps  $(T_i - T_{i-1})_{i \geq 1}$  sont des v.a.i.i.d. de loi  $\mathcal{E}(\lambda)$ . Les durées d'envoi  $E_1, E_2, \dots$  ( $E_i$  est la durée d'envoi du  $i^{\text{ème}}$  message) des messages sont des v.a.i.i.d. indépendantes des instants d'arrivées de loi  $\mathcal{E}(\mu)$ , où  $\mu > 0$ . On note  $X_t$  l'encombrement de la mémoire tampon (message en cours d'envoi inclus) à l'instant  $t$ . On a donc

$$X_t = \begin{cases} 0 & \text{si } t < T_1 \\ 1 & \text{si } T_1 \leq t < \min\{T_2, T_1 + E_1\} \\ 0 & \text{si } T_1 + E_1 < T_2 \text{ et } T_1 + E_1 \leq t < T_2 \\ 2 & \text{si } T_1 + E_1 > T_2 \text{ et } T_2 \leq t < \min\{T_1 + E_1, T_3\} \\ \vdots & \end{cases}$$

Le processus  $(X_t)_{t \in \mathbb{R}^+}$  est connu dans la littérature sous le nom de file d'attente de type M/M/1 (voir par exemple [Bou02] ou même [Wik]). Il pourra être simulé, selon le contexte, de façon exacte ou en discrétisant le temps. Afin de discrétiser le temps, il convient de bien comprendre l'évolution du processus  $(X_t)$  : il découle de la propriété d'*absence de mémoire* des lois exponentielles que connaissant la valeur de  $X_t$  à un certain instant  $t = t_0$ , on peut simuler son devenir  $(X_t)_{t \in [t_0, +\infty[}$  de la façon suivante :

**1er cas :**  $X_{t_0} = 0$  (aucun message en attente d'envoi) : on tire une variable aléatoire  $T$  de loi  $\mathcal{E}(\lambda)$ ,  $X_t$  garde la valeur 0 pour  $t_0 \leq t < t_0 + T$ , et prend la valeur 1 en  $t = t_0 + T$ .

**2ème cas :**  $X_{t_0} \geq 1$  (au moins un message en attente d'envoi) : on tire deux variables aléatoires  $T$  et  $E$ , indépendantes et de lois respectives  $\mathcal{E}(\lambda)$  et  $\mathcal{E}(\mu)$ ,  $X_t$  garde la valeur de  $X_{t_0}$  pour  $t_0 \leq t < t_0 + \min\{T, E\}$ , et, en  $t = t_0 + \min\{T, E\}$ ,  $X_t$  prend la valeur  $X_{t_0} + 1$  si  $T < E$  et la valeur  $X_{t_0} - 1$  si  $T \geq E$ . Mais il est simple de voir que la loi de  $\min\{T, E\}$  est  $\mathcal{E}(\lambda + \mu)$  et que, connaissant la valeur de  $\mathcal{E}(\lambda + \mu)$ , la probabilité que ce minimum soit  $T$  (resp.  $E$ ) est  $\frac{\lambda}{\lambda + \mu}$  (resp.  $\frac{\mu}{\lambda + \mu}$ ) : on peut donc ne tirer qu'une variable aléatoire  $Y$  de loi  $\mathcal{E}(\lambda + \mu)$  ainsi qu'une variable aléatoire  $\varepsilon$  qui vaut 1 ou  $-1$  avec probabilités respectives  $\frac{\lambda}{\lambda + \mu}$  et  $\frac{\mu}{\lambda + \mu}$  et laisser  $X_t$  garder la valeur de  $X_{t_0}$  pour  $t_0 \leq t < t_0 + Y$ , et, en  $t = t_0 + Y$ , poser  $X_t = X_{t_0} + \varepsilon$ .

On réitère ensuite le même procédé pour déterminer l'évolution de  $X_t$  au delà.

## ● Objectifs

Il s'agit tout d'abord de comprendre comment le processus  $X_t$  évolue (tendance à la stabilité ou non, valeur moyenne) en fonction des paramètres  $\lambda$  et  $\mu$ . Au delà de ce préliminaire, on choisit  $\lambda < \mu$  et, pour une taille  $N \gg 1$  de la mémoire tampon et un seuil de temps  $S \gg 1$ , on essaye d'estimer l'espérance du *temps de saturation* :

$$T_N := \min\{t \geq 0; X_t = N\}$$

et la probabilité de l'événement de *saturation avant*  $S$  :

$$T_N \leq S.$$

On dimensionnera ensuite la taille  $N$  de la mémoire tampon de façon à ce que

$$\mathbb{E}[T_N] > S$$

ou

$$\mathbb{P}(T_N \leq S) \leq \alpha$$

avec  $S$  un seuil élevé et  $\alpha$  un niveau de risque faible.

## ● Feuille de route

a) Les deux premiers enjeux de ce projet sont :

1. La détermination d'une condition sur  $\lambda$  et  $\mu$  pour la *stabilité* du processus (*i.e.* pour que la valeur de  $X_t$  n'ait pas tendance à exploser lorsque  $t$  devient grand).

2. La maîtrise de la *discrétisation en temps* du processus. Il va être en effet profitable de le remplacer par une chaîne de Markov à temps discret, mais il sera nécessaire de calibrer le pas de temps choisi de façon à ce que le processus à temps discret garde le même comportement que le processus continu (ce qui impose un pas de temps relativement petit), sans impliquer trop de calculs (ce qui exclut un pas de temps excessivement petit).

On pourra donc commencer par simuler le processus en temps continu et en temps discret et conjecturer un critère de stabilité, pour lequel on pourra aussi se documenter *via* [Bou02] ou [Wik]. On calculera  $\mathbb{E}[T_N]$  et  $\mathbb{P}(T_N \leq S)$  pour des valeurs raisonnables de  $N$  et  $S$  et on en déduira une calibration de la discrétisation du temps de façon à ce que le modèle discret soit proche du modèle continu. On rappelle que l'analogue discret d'une loi exponentielle est la loi géométrique de même espérance. Pour des exemples de valeurs numériques pour  $\lambda$  et  $\mu$ , on pourra consulter la Table II de l'article [PW89].

b) De façon à aborder le cas où  $N$  est relativement élevé, faire un changement de probabilité pour le modèle discret en modifiant de façon intuitive les paramètres  $\lambda$  et  $\mu$ , estimer  $\mathbb{E}[T_N]$  et  $\mathbb{P}(T_N \leq S)$  (on pourra éventuellement s'inspirer de la partie II de [PW89]).

c) Toujours pour le modèle discret, afin d'estimer  $\mathbb{E}[T_N]$  et  $\mathbb{P}(T_N \leq S)$  pour  $N$  élevé, optimiser son changement de probabilité.

## ● Extensions

Pour la suite du projet, au moins une direction parmi les suivantes devra être prise :

d) On peut modifier le modèle de façon à tenir compte de la *taille* des paquets :  $X_t$  est alors la somme des tailles des paquets arrivés et non envoyés jusqu'à l'instant  $t$ . La taille d'un paquet sera une variable aléatoire, éventuellement dépendante de sa durée d'envoi. Il existe de multiples façons de générer de la dépendance entre deux v.a.  $T$  et  $D$ . En voici quelques exemples :

- on peut choisir  $D$  comme étant une fonction de  $T$ ,
- on peut choisir  $D$  comme étant une fonction de  $T$  et d'une autre variables aléatoire indépendante de  $T$  (par exemple  $D = \alpha T + (1 - \alpha)Y$ , avec  $\alpha \in ]0, 1[$  et  $c > 0$  des coefficients et  $Y$  une v.a. indépendante de  $T$  de loi exponentielle),
- On peut choisir  $D$  comme étant distribuée selon une loi exponentielle dont le coefficient est une fonction de la valeur de  $T$ ,
- ...

d') Supposons maintenant que, avant d'être envoyés, les messages doivent être traités (par exemple compressés) : avant son envoi, chaque message arrivant fait l'objet d'un traitement nécessitant un temps de loi  $\mathcal{E}(\tau)$  indépendant

des autres aléas et ne pouvant être réalisé sur deux paquets en même temps. En s'inspirant de la théorie des réseaux de Jackson à guichets multiples (voir par exemple le dernier chapitre de [Bou02]), on essayera de mettre en évidence un critère de stabilité (pour l'ensemble des deux files d'attente) et de calibrer la taille de la mémoire tampon.

d") On pourra essayer d'utiliser la théorie des grandes déviations (voir [Buc04] ou [SW95]) pour proposer un autre échantillonnage préférentiel. On pourra notamment utiliser Theorem 11.3 dans [SW95] pour identifier la trajectoire typique menant  $X_t$  à l'état  $N \gg 1$ .

## ● Références bibliographiques

Sur les files d'attente et les processus de sauts : [Bou02], [Wik].

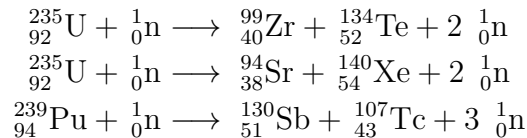
Sur leur simulation : [PW89].

Sur les grandes déviations : [Buc04], [SW95].

## 2 Sûreté d'un réacteur nucléaire

Le milieu du XX<sup>e</sup> siècle a marqué le début du développement du nucléaire aux Etats-Unis, en URSS, en France ... voir <http://nucleaire.cea.fr/fr/repere/historique.htm>. En dehors des applications militaires (bombe atomique, propulsion nucléaire navale), les applications civiles concernent surtout la production d'électricité via les centrales nucléaires. Depuis les années 1960, les technologies ont considérablement évolué, utilisant surtout le principe de fission alors que les projets à l'étude concernent la fusion (ITER par exemple).

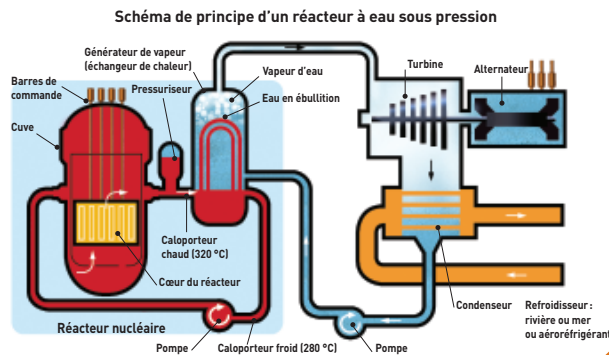
La fission nucléaire, découverte à la fin des années 30, est le phénomène de fragmentation de noyaux lourds (Uranium 235 ou 238, Plutonium 239) par capture de neutrons, donnant lieu à des nucléides plus légers et produisant une énergie importante. Par exemple



La fission produit de nouveaux neutrons qui peuvent entraîner une réaction en chaîne. L'énergie thermique dégagée est ensuite transformée (production de vapeur) en énergie mécanique ou électrique. Pour permettre à la réaction de continuer, il faut une quantité suffisante de combustible nucléaire dans le réacteur. A l'inverse, pour éviter un emballement de la réaction, des barres de commande peuvent être introduites dans le cœur du réacteur : ces barres constituées de matériaux spéciaux ont pour effet d'absorber les neutrons et ainsi d'arrêter la réaction.

L'accident de Three Mile Island en 1979 et la catastrophe de Tchernobyl en 1986 avaient déjà rappelé les risques du nucléaire. La catastrophe de Fukushima en 2011 a relancé le débat sur l'évaluation des risques extrêmes, dans un contexte où nos sociétés s'interrogent sur nos modèles de dépendance énergétique<sup>2</sup> et le développement d'énergies renouvelables.

Dans une centrale nucléaire, le danger le plus important provient d'un défaut de refroidissement du cœur du réacteur (avec pour conséquence un risque



Source : <http://www.cea.fr/content/download/4687/27890/file/livret-reacteur.pdf>

2. En France, la filière nucléaire fournit à elle seule 78% de la production électrique.

de fusion libérant les éléments très radioactifs). La prise en compte de tous les facteurs est complexe<sup>3</sup> et nous étudions un modèle simplifié, cherchant à évaluer la capacité des barres de commande à absorber les neutrons (arrêt d'urgence de la réaction), dans un contexte où les réflecteurs entourant le cœur sont endommagés (fissure, dégradation des propriétés réfléchissantes).

## ● Modèle neutronique

Nous modélisons le déplacement d'un seul neutron. Pendant son parcours (modélisé dans cette partie dans le plan), il est soumis à de nombreux chocs qui se produisent de manière imprévisible : à chaque choc, sa vitesse saute d'une quantité tirée aléatoirement et indépendamment de son parcours précédent. Si  $(\tau_i)_i$  est la suite des instants de chocs, sa vitesse (en km/s) est

$$V_t = V_{\tau_i^-} + Y_i, \quad \text{pour } t \in [\tau_i, \tau_{i+1})$$

où  $(Y_i)_i$  est une suite de v.a.i.i.d. dans  $\mathbb{R}^2$  modélisant les sauts de vitesse (chocs). La vitesse à la date  $t$  est donc :

$$V_t = V_0 + \sum_{i=1}^{N_t} Y_i$$

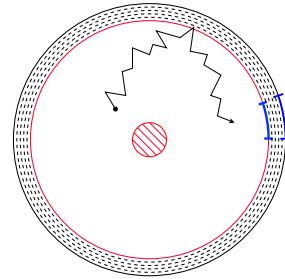
où  $N$  est un processus de Poisson d'intensité  $\lambda$  grande. Les variables  $(Y_i)_i$  et le processus  $N$  sont indépendants. La loi de  $Y$  est centrée et on pourra choisir une loi gaussienne sur chaque coordonnée, d'écart type  $2000 \text{ km s}^{-1}$  (neutron lent). Les ordres de grandeur réalistes pour  $\lambda$  sont de l'ordre  $10^{12} \text{ s}^{-1}$  à  $10^{15} \text{ s}^{-1}$ .

Si l'on ne tient pas compte des barres de commande et de l'enceinte du réacteur, la position du neutron au cours du temps évolue donc comme

$$X_t = X_0 + \int_0^t V_s ds.$$

C'est un modèle simple d'*équation de transport*, voir [LPS98]. Le cœur du réacteur est composé d'une enceinte extérieure réfléchissante (qu'on suppose circulaire de diamètre  $D = 5m$ ), et au centre d'une barre de commande (de diamètre  $d$ ). Plusieurs situations peuvent se présenter.

- Lorsque le neutron touche la barre de commande, il est absorbé et s'arrête.



Modèle de réacteur et déplacement de neutron

3. voir la conférence [http://www.coriolis.polytechnique.fr/Confs/Herviou\\_conf.pdf](http://www.coriolis.polytechnique.fr/Confs/Herviou_conf.pdf)



- Lorsque le neutron atteint l'enceinte réfléchissante, il est réfléchi vers l'intérieur.
- Une partie de l'enceinte est endommagée, sur une portion angulaire mesurant quelques mm. On peut imaginer que si le neutron atteint cette partie, il quitte le cœur du réacteur.

On cherche à évaluer la probabilité de la perte de contrôle du neutron décrit par l'évènement

$A = \{\text{le neutron atteint la zone endommagée avant d'être absorbée par la barre}\}.$

▷ QUESTIONS À TRAITER :

- Approximation asymptotique du modèle lorsque l'intensité est très grande et application.
- Evaluer  $\mathbb{P}(A)$  en fonction des positions et vitesses initiales.
- Déterminer la taille de la barre de commande pour maintenir une probabilité inférieure à  $10^{-5}$ .
- Déterminer la distribution du temps d'absorption conditionnellement à  $A^c$  ou du temps de sortie conditionnellement à  $A$ .

## ● Réacteur à plusieurs barres de commande

Le cœur du réacteur contient maintenant 4 barres de commande disposées en carré. A quelle distance du centre disposer les barres pour garantir une efficacité maximale ? Discuter de l'influence des autres paramètres.

## ● Références bibliographiques

Sur le nucléaire : <http://www.cea.fr/content/download/4687/27890/file/livret-reacteur.pdf>.

Site de l'Autorité de Sûreté Nucléaire : <http://www.asn.fr/index.php>.

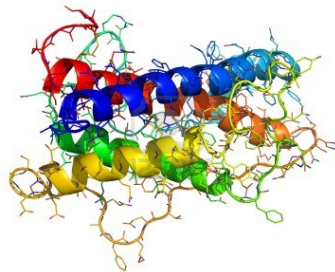
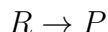
Sur les équations de transport : [LPS98] ou [CDL<sup>+</sup>89].

Sur la neutronique : [BR78].

Sur les asymptotiques de diffusion : [Pap75].

### 3 Réaction chimique

La simulation aléatoire occupe un grand rôle en chimie (ou biochimie) où elle est couramment utilisée notamment pour explorer des conformations de moindre énergie (comme par exemple le repliement des protéines ; voir aussi [Pot]). On se propose de modéliser au niveau microscopique le taux d’une réaction chimique. En effet la transformation d’un réactif  $R$  en un produit  $P$ , écrite



peut se voir comme la transition d’une molécule (ou d’un ensemble de molécules) de l’état  $R$  dans un état  $P$ . On suppose ces différents états caractérisés par une énergie potentielle  $E$ . Cette énergie dépend *a priori* d’un vecteur  $x \in \mathbb{R}^n$  décrivant la position des éléments (distances entre atomes, angles relatifs...). Le modèle mis en œuvre (qui est purement classique et ne tient aucun compte des effets quantiques) est le suivant [Kra40] : chaque degré de liberté de la molécule subit une agitation thermique (qui va être modélisée par un mouvement aléatoire) et subit l’énergie potentielle.

▷ **Modélisation.** On se donne un pas de temps  $h > 0$ , et l’évolution de ce vecteur  $(X_{ih})_{i \geq 0}$  est décrite par

$$X_{ih+h} = X_{ih} - \nabla E(X_{ih})h + \sqrt{2\varepsilon h}W_i,$$

où  $\varepsilon > 0$  est un paramètre petit (identifié à la température car il décrit l’agitation thermique) et  $(W_i)$  est une suite de vecteurs aléatoires i.i.d. à composante indépendante de loi normale centrée réduite.

Dans les situations réelles, il est délicat de calculer l’énergie  $E$  ; des méthodes existent (voir [Pot] et [SJ96] notamment sur l’énergie d’une liaison alcane  $C - C - C - C$ ). On suppose que l’énergie  $E$  ne dépend que d’un seul paramètre (coordonnée de réaction), le processus  $X$  vit donc sur  $\mathbb{R}$ , il est unidimensionnel. Les états  $R$  et  $P$  correspondent à des minima locaux de  $E$ .

La géométrie de l’énergie  $E$  est alors très importante : en particulier, on va supposer que  $E$  n’a que ces deux minima ; entre  $R$  et  $P$  il y a alors un unique maximum  $T$  (comme état de transition). On prendra dans un premier temps

$$E(x) = \frac{x^4}{4} - \frac{x^2}{2} + \frac{x^3}{3}. \quad (1)$$

On peut avoir deux cas, soit  $E(R) > E(P)$  soit  $E(R) < E(P)$ .

## • Objectifs

▷ **Temps de transition.** Pour commencer, on étudie une seule molécule, et notamment son temps de transition  $\tau$  nécessaire pour passer de l'état  $R$  à l'état  $P$ , ou vice-versa. On suppose que  $E(R) > E(P)$  et  $E''(T) < 0$ . On prend  $\varepsilon < \frac{1}{2}(E(P) - E(R))$ . On s'intéresse à la probabilité  $\mathbb{P}(\tau < t_0)$  qu'une transition ait lieu avant le temps  $t_0$ , pour  $t_0$  fixé.

1. Donner les états  $P$ ,  $R$  et  $T$  pour l'énergie  $E$  (1).
2. Calculer numériquement la probabilité  $\mathbb{P}(\tau < t_0)$  pour une molécule réagissant de  $R$  à  $P$ , ou de  $P$  vers  $R$ . Expliquer de manière heuristique la différence entre les deux valeurs.
3. Quand  $\varepsilon$  tend vers 0, la variable aléatoire  $\frac{\tau}{\mathbb{E}[\tau]}$  semble-t-elle converger en loi ? Identifier la loi.

On fera attention à l'influence de  $h$  sur les résultats.

▷ **Taux de réaction.** On va maintenant s'intéresser à un ensemble important de particules. On considère que les particules sont dans l'état  $R$  ou  $P$  si elles sont dans le bassin d'attraction de l'énergie correspondant. Dans le cas (1), la particule  $X_t^1$  est dans l'état  $R$  si  $X_t^1 > 0$  et dans l'état  $P$  si  $X_t^1 < 0$ . Dans d'autres cas de potentiel, on pourra choisir de petits voisinages centrés autour des minima. On appelle le taux de réaction, la valeur  $\kappa$  définie empiriquement par

$$\frac{d[P]}{dt} = \kappa[R] \quad (2)$$

où  $[A]$  désigne la concentration de l'état  $A$ .

1. Calculer numériquement le taux de réaction  $\kappa$  ; comment évolue-t-il avec  $\varepsilon$  ? Comparer avec  $\frac{1}{\mathbb{E}[\tau]}$ . Expliquer heuristiquement l'équation (2) en fonction de la question 3 précédente.
2. On pourra comparer la valeur obtenue de  $\mathbb{E}[\tau]$  (ou de  $\kappa$ ) avec la valeur suivante (formule d'Eyring-Kramers [BEGK04, HTB90, Kra40])

$$T_0 = \frac{2\pi}{\sqrt{|E''(T)| E''(R)}} e^{(E(T) - E(R))/\varepsilon}$$

et justifier numériquement que ces valeurs coïncident lorsque  $\varepsilon$  tend vers 0.

## • Extensions : impact du potentiel

Le potentiel joue un rôle important dans ces problèmes. On cherchera à traiter au moins l'une des extensions suivantes.

- a) Si  $E$  ne vérifie pas les conditions précédentes, notamment si  $E$  est quartique en  $T$  i.e  $E''(T) = 0, E^{(4)}(T) < 0$  ou si  $E$  non dérivable en  $T$ , analyser empiriquement comment évolue la vitesse de réaction en fonction de  $\varepsilon$ . La réaction a-t'elle lieu plus vite ou moins vite ?
- b) Que se passe-t-il si  $E$  possède un intermédiaire de réaction supplémentaire (voir par exemple figure 1) ? On pourra regarder plusieurs cas

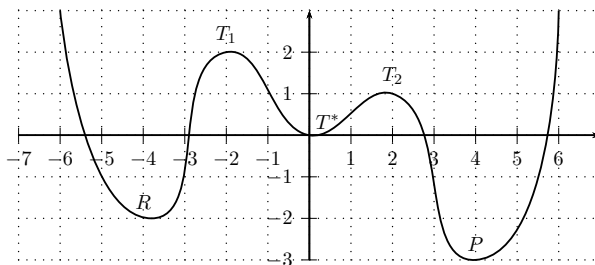


FIGURE 1 – Exemple de potentiel à trois puits :  $T^*$  produit intermédiaire de réaction.

dont notamment  $E(T_1) = E(T_2)$ . La vitesse de réaction est-elle plus ou moins favorisée par l'apparition de ce produit intermédiaire ? Regarder aussi la stabilité du produit  $T^*$  : est-il plus ou moins stable ? à quelle vitesse se transforme-t'il ? vers quels produits ? On pourra pour cela estimer la probabilité  $\mathbb{P}_{T^*}[\tau(R) < \tau(P)]$ .

Enfin étendre l'analyse en dimension 2 avec un potentiel bi-dimensionnel du type de la figure 2, c'est-à-dire  $E(x, y) = \frac{x^4 + y^4}{4} - \frac{x^2 + y^2}{2} + \gamma \frac{(x - y)^2}{2}$  avec  $\gamma < 1/2$ .

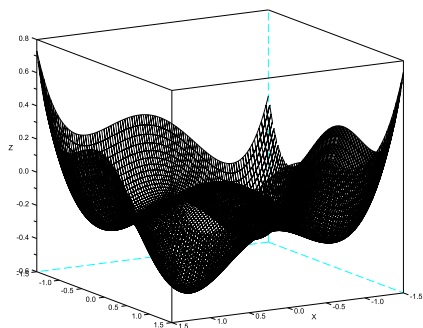


FIGURE 2 – Exemple de potentiel.

## ● Références bibliographiques

Sur des exemples en chimie : [Pot], [SJ96].

Sur le calcul du temps d'atteinte : [OV05, Chap. 5], [BEGK04].

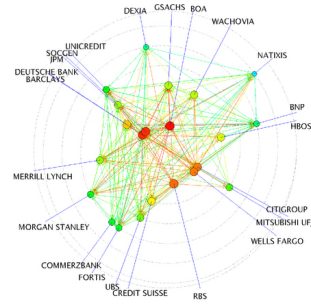
Sur l'historique : [Kra40], [HTB90].

## 4 Risque systémique

### ● Réseau financier et risque systémique d'un défaut

La crise financière que nous traversons est un évènement d'une rare ampleur et lourd en conséquence. Elle est qualifiée de systémique car elle s'est manifestée à l'échelle du système bancaire mondiale. Toutes les banques ont du faire face à de grandes difficultés de liquidités ayant conduit certaines d'entre elles à la faillite (Lehman Brothers en 2008). Depuis le début de la crise financière en 2007, plus de 370 des banques américaines (sur près de 8000 banques assurées par la Federal Deposit Insurance Corporation) ont fait faillite. Entre 2000 et 2004, seulement 30 banques ont fait faillite et aucune entre 2005 et le début de l'année 2007.

Le risque systémique est le risque d'effondrement d'un système suite à un choc sur certaines institutions financières qui entraînent par un effet *domino* la dégradation brutale ou la faillite de beaucoup d'autres. Ce n'est que très récemment que les institutions financières se sont intéressées à la modélisation mathématique de ces épisodes de contagion par défaut où un choc économique causant des pertes initiales et le défaut de quelques institutions sont amplifiés en raison de liens financiers complexes, pour finalement conduire à des faillites à plus grande échelle.



Risque de contagion entre banques au cœur de la crise (d'après <http://www.nature.com/srep/2012/120802/srep00541/full/srep00541.html>)

### ● Modélisation du bilan d'une banque et réseaux financiers

Un système d'institutions financières est naturellement modélisé par un réseau de relations ou contreparties (graphe) : un ensemble de  $n$  noeuds et des liens pondérés représentés par une matrice  $E = (e_{i,j})_{1 \leq i,j \leq n}$  où  $e_{i,j}$  est (la valeur marché de) l'exposition de l'institution financière  $i$  à l'institution financière  $j$ .

A la date  $t$ , l'institution  $i$  dispose d'un capital (propre)  $X_i(t)$ , c'est-à-dire un matelas de sécurité pour les créanciers de l'entreprise, pour absorber les pertes potentielles. Le bilan d'une banque se présente sous la forme d'un équilibre entre ses actifs<sup>4</sup>, c'est-à-dire les biens qu'elle possède qui ont une valeur économique positive, et ses passifs, c'est-à-dire l'ensemble de ses dettes ou de ses biens qui ont une valeur économique négative, c.f. Table 1.

4. on ne considère ici que des actifs interbancaires :  $A_i(t) \equiv 0$ .

Actifs	Passifs
Actifs interbancaires $\sum_{j=1}^n e_{i,j}$	Passifs interbancaires $\sum_{j=1}^n e_{j,i}$
Autres actifs $A_i(t) = 0$	Capital $X_i(t)$

TABLE 1 – Bilan Actifs-Passifs de l’institution financière  $i$  à la date  $t$

Entre les temps  $t_k$  et  $t_{k+1}$  ( $t_k := k\delta$  avec  $\delta > 0$ ), le capital d’une institution subit des fluctuations dues au marché dont la dynamique peut s’écrire

$$X(t_{k+1}) = e^{-\lambda\delta}X(t_k) + \mu(1 - e^{-\lambda\delta}) + \sigma\sqrt{\delta}W_k \quad (3)$$

où  $(W_k)_{k \geq 0}$  est une suite de v.a.i.i.d. gaussiennes centrées réduites. Le paramètre  $\mu$  s’interprète comme l’équilibre moyen du capital, c’est-à-dire que lorsque  $X(t)$  s’écarte de cette valeur le paramètre  $\lambda$  agit comme une force de rappel vers  $\mu$ , et le paramètre  $\sigma > 0$  est l’écart-type des fluctuations du capital. Le processus  $(X_{t_k})_{k \geq 0}$  est alors un processus gaussien.

▷ **Bilan à l’horizon  $T$ .** A la date  $T = N\delta$ , un bilan de solvabilité a lieu permettant de conclure quant à la solidité financière de l’institution  $i$ . Si le capital de l’institution  $i$  est en dessous d’un seuil critique déterministe  $c_i$  celle-ci n’est plus *solvable*. Une banque est solvable si son capital restant à la date  $T$  est supérieur au seuil critique, c’est-à-dire si  $X_i(T) > c_i$ . Une banque insolvable fait défaut et est liquidée. Chacun de ses créanciers perd une fraction  $1 - R$  de l’exposition à la banque faisant défaut. Cette perte vient alors se soustraire au capital et peut entraîner à son tour l’insolvabilité des créanciers. Cette cascade de défaut dépend fortement du taux de récupération  $R$  (recovery rate) de la banque faisant défaut : pour simplifier, on suppose que ce taux de récupération est le même pour toutes les institutions (on prendra par exemple  $R = 5\%$ ).

On définit l’ensemble  $D_0^T = \{i \in \{1, \dots, n\} : X_i(T) < c_i\}$  des institutions financières faisant initialement défaut dans le réseau à la date  $T$ , la cascade de défaut est la séquence d’ensemble  $D_0^T \subset D_1^T \subset \dots \subset D_{n-1}^T$ , définie par

$$D_k^T = D_{k-1}^T \cup \left\{ j \notin D_{k-1}^T : X_j(T) - \sum_{p \in D_{k-1}^T} (1 - R)e_{j,p} < c_j \right\}, \quad k \geq 1.$$

Dans un réseau financier de taille  $n$ , cette cascade finit après au plus  $n - 1$  étapes. A l’étape  $k$ ,  $D_k^T$  représente l’ensemble des institutions financières insolubles (et donc faisant défaut) suite à l’exposition de contrepartie à des banques de l’ensemble  $D_{k-1}^T$  qui viennent de faire défaut à l’étape précédente.

Pour quantifier le risque systémique et l'effet de contagion, on définit l'impact de défaut  $I(T)$  à la date  $T$  du à la cascade de défaut à l'instant  $T$ ,

$$I(T) = \sum_{j \in D_{n-1}^T} \left( X_j(T) + \sum_{p \notin D_{n-1}^T} (1 - R) e_{p,j} \right)$$

qui est la somme des pertes générées par la contagion du défaut des banques à la date  $T$ .

▷ **Capitaux indépendants.** On supposera dans un premier temps que les capitaux évoluent de manière indépendante : les suites de variables  $(W_k)_{k \geq 0}$  apparaissant dans (3) pour chaque institution sont mutuellement indépendantes.

L'horizon  $T$  est 1 an et les dates  $(t_k)_k$  d'observation d'évolution du capital sont trimestrielles ou mensuelles par exemple. Les autres paramètres sont  $X_i(0) = \mu = 15$ ,  $\sigma = 8$ ,  $\lambda = 20$ ,  $c_i = 10$ ,  $i = 1, \dots, 5$  avec un réseau simple constitué de 5 institutions financières, dont les liens interbancaires pourront être donnés par la matrice

$$E = \begin{pmatrix} 0 & 3 & 0 & 0 & 6 \\ 3 & 0 & 0 & 0 & 0 \\ 3 & 3 & 0 & 0 & 0 \\ 2 & 2 & 2 & 0 & 2 \\ 0 & 2 & 3 & 3 & 0 \end{pmatrix}.$$

On abordera les points suivants.

1. Estimation de la probabilité que le nombre de banques insolubles soit 1, 2, 3, 4 ou 5.
2. Estimation de la Value-at-Risk et de la Conditional Value-at-Risk de la distribution  $I(T)$  pour différents seuils  $\alpha$  proches de 1 (par exemple  $\alpha = 99.9999\%$ ), ainsi que des intervalles de confiance associés aux estimateurs.
3. Estimation de la cascade d'insolvabilité la plus probable, selon le nombre de banques en défaut.
4. Identification des liaisons de contrepartie dangereuses, des scénarios critiques menant à l'effondrement du système ainsi que des configurations (raisonnables) permettant d'améliorer la sûreté du système.
5. Influence de la taille du réseau ( $n = 10, 50$ ) et des configurations dangereuses.
6. Evaluation de la distribution de l'impact de défaut conditionnellement à l'évènement : "le réseau entier d'institutions financières s'est effondré".

## ● Extensions

On cherchera à aborder l'un des points suivants.

▷ **Contagion dynamique.** En réalité, l'insolvabilité des institutions peut avoir lieu à tout moment entre la date 0 et la date  $T$ . Il est donc nécessaire de prendre en compte ce risque de manière dynamique. Cela nous conduit à considérer des impacts de défaut dynamique  $I(t_k)$ ,  $k = 1, \dots, N$ . On définit l'impact de défaut total à la date  $T$  comme la somme de tous les impacts de défaut entre la date 0 et la date  $T$

$$\sum_{k=1}^N I(t_k).$$

On reprendra les questions précédentes en intégrant le risque de contagion dynamique.

▷ **Modèle d'équilibre long terme.** Un modèle couramment utilisé pour les simulations de risque de défaut est un modèle d'équilibre à 1 facteur idiosyncratique et 1 facteur systémique. Les capitaux de toutes les institutions ont un facteur commun  $(Z_{t_k})_{k \geq 0}$  régissant l'équilibre long terme du réseau financier. Ce capital long terme est un processus gaussien de dynamique

$$Z(t_{k+1}) = e^{-\lambda_e \delta} Z(t_k) + \sigma_e \sqrt{\delta} \bar{W}_k,$$

où  $(\bar{W}_k)_{k \geq 0}$  est une nouvelle suite indépendante de v.a.i.i.d. gaussiennes centrées réduites et  $\lambda_e = 10$ ,  $\sigma_e = 3$ . Ce capital d'équilibre vient s'ajouter à chaque capital  $X_i(t)$  de l'institution  $i$ . On reprendra les questions précédentes avec ce nouveau modèle.

## ● Références bibliographiques

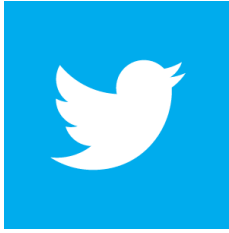
Sur le risque systémique : [Min11], [Mou11].

Sur le calcul des risques extrêmes par simulation : [GHS00], [Gla03].

Sur les mesures de risques : [FS02].



## 5 Messages twitter



Twitter est une plateforme de microblogage, permettant d'envoyer rapidement de brefs messages (pas plus de 140 caractères), appelés **tweets** ("gazouillis"), sur internet. Cette plateforme est devenue très populaire, et a atteint les 500 millions d'utilisateurs en février 2012.

Ces messages contiennent (parmi d'autres) plusieurs informations : le moment de l'envoi du message (sous la forme d'un timestamp UNIX <sup>5</sup>), l'auteur et le corps du message. Certains messages forment des conversations, d'autres sont des messages décrivant un événement venant d'arriver dans le monde. Le nombre très élevé d'utilisateurs de twitter en fait une source d'information en temps réel très intéressante.

De nombreux tweets sont par exemple envoyés par des acteurs des marchés financiers. Quelques articles prétendent même que twitter permettrait de mieux prédire le comportement d'un marché financier que des analyses plus classiques <sup>6 7</sup>. Un exemple de tweet est le suivant :

```
User:  SchwartzNow
Time:  2009-12-17T19:20:15
Body:  also for tomorrow expect high volume options
       traded stocks like $aapl,$goog graviate around
       the strikes due to the delta hedging
```

Certains mots portent le sens du message : **tomorrow, high volume, options, stocks, \$aapl, \$goog**. Ce tweet contient des mots commençant par le symbole \$, appelés des "cashtags" : \$aapl et \$goog (AAPL et GOOG sont les tickers <sup>8</sup> des actions Apple et Google). Un exemple d'un tel jeu de données est disponible ici <sup>9</sup>.

### ● Objectifs

On veut construire un modèle simple pour le flux de tweets observés sur une période de temps. Les phénomènes observés dans ce type de flux, que l'on souhaite incorporer au modèle, sont les suivants :

---

5. Le timestamp UNIX correspond, par convention, au nombre de millisecondes écoulées depuis le 1er janvier 1970.

6. <http://www.dailymail.co.uk/sciencetech/article-2120416/Twitter-predicts-stock-prices-accurately-investment-tactic-say-scientists.html>

7. <http://singularityhub.com/2010/10/21/twitter-predicts-the-stock-market/>

8. abréviation utilisée pour désigner une action sur les marchés financiers

9. <http://www.infochimps.com/datasets/twitter-census-stock-tweets>

1. La plupart des tweets sont envoyés à cause d'un événement observé dans le monde, sans lien avec d'autres tweets.
2. Un tweet peut générer des nouveaux tweets (une réponse à ce tweet, ou un "retweet"<sup>10</sup>). Certains tweets contiennent une information suscitant beaucoup de réactions. Cela génère un grand nombre de tweets, qui eux-même génèrent un grand nombre de tweet, etc. Dans ce flux, on observe donc une "auto-excitation", qui peut produire des périodes de temps contenant beaucoup de tweets.

L'objectif est d'évaluer la probabilité d'avoir un nombre très élevé de tweets dans une période de temps donnée, en fonction du régime d'auto-excitation considéré, et de la durée de cette période. On cherchera notamment à intégrer dans la modélisation le fait qu'un tweet a un impact sur le flux qui diffère selon son contenu. Les points à traiter sont les suivants.



Tweets (en rose) et retweets (en vert) dans l'heure qui a suivi le tremblement de terre au Japon (mars 2011).  
SOURCE : <http://www.flickr.com/photos/twitteroffice/5884626815/>

1. Simuler le flux de tweet sur une journée, en prenant la seconde comme unité de temps. On paramétrisera le modèle pour avoir, en dehors du phénomène d'auto-excitation, 1 tweet par seconde. Proposer une paramétrisation du modèle permettant de modéliser une auto-excitation faible ou forte, qui persiste longtemps ou diminue rapidement dans le temps.
2. La forte croissance de Twitter a conduit il y a quelques années à la "twitpocalypse"<sup>11 12</sup>, qui est survenue quand le nombre d'identifiants possibles des tweets (qui étaient alors codés avec un entier sur 32 bits) a été dépassé. Evaluer la probabilité d'une twitpocalypse en fonction de la période de temps (jour, semaine, mois), et du niveau d'auto-excitation du modèle.
3. Chaque tweet a des caractéristiques ("features") qui ont un impact plus ou moins fort sur l'auto-excitation du flux. Des exemples de features sont la présence d'un mot, ou d'une combinaison de mots dans le tweet, la présence d'un URL, etc. Proposer une modélisation simple de ce phénomène à l'aide d'une dizaine de features possibles.

Simuler des configurations avec des features qui ont une très faible probabilité d'apparaître dans un tweet (de l'ordre de 0.01%, mais qui ont

10. un utilisateur fait un copier/coller du tweet et le renvoie

11. <http://www.twitpocalypse.com>

12. <http://techcrunch.com/2009/06/12/all-hell-may-break-loose-on-twitter-in-2-hours/>

un impact très fort sur le flux). Simuler également des configurations avec des *features* récurrentes (probabilité de l'ordre de 1%), mais qui ont un faible impact sur le flux.

Déterminer le nombre moyen de tweets apparaissant dans l'heure (ou plus) suivant un tweet contenant une feature avec un fort impact. Evaluer la probabilité d'une twitpocalypse sur différentes périodes de temps, conditionnellement au fait qu'un message contienne une feature avec un fort impact.

## ● Une modélisation de l'auto-excitation (sans impact de la feature)

On se place dans un intervalle de temps  $[0, T]$ , l'unité de temps est la seconde, et  $T$  correspond à une journée. On note  $t_1 < t_2 < \dots < t_{N_T}$  la suite des temps d'envoi d'un tweet, tombant dans l'intervalle  $[0, T]$ , où on a posé  $N_t = \sum_{i \geq 1} \mathbf{1}_{\{t_i \leq t\}}$ , le processus de comptage associé à cette suite. On modélise la suite  $\{t_i\}_{i \geq 1}$  par un processus ponctuel "auto-excité", appelé aussi processus de Hawkes [Haw71], en s'inspirant de [SJ12]. Ce type de processus est également très utilisé pour modéliser les tremblements de terre [Oga98]. Dans ce type de modèle, l'intensité du processus de comptage  $\{N_t : t \geq 0\}$  est de la forme

$$\lambda(t) = \lambda_0 + \sum_{i \geq 1, t_i \leq t} g(t - t_i), \quad (4)$$

où  $\lambda_0 > 0$  et  $g$  est une fonction déterministe  $\mathbb{R} \rightarrow \mathbb{R}^+$  telle que  $g(t) = 0$  pour tout  $t \leq 0$ , et décroissante sur  $\mathbb{R}^+$ . L'intensité  $\lambda$  est donc aléatoire, et  $\lambda(t)$  dépend, à l'instant  $t$ , des temps  $t_i < t$ . Cela garantit la propriété de "causalité" : seuls les tweets passés ont une influence éventuelle sur les tweets à venir. La décroissance de  $g$  modélise le fait que l'impact d'un tweet décroît avec le temps. La forme de cette fonction quantifie la "vitesse" de cette décroissance.

L'intensité (4) est la somme des fonctions d'intensité  $\lambda_0, g(\cdot - t_1), g(\cdot - t_2), \dots$ . On peut donc simuler un processus ponctuel d'intensité (4) en "superposant" des générations de processus de Poisson admettant ces intensités. Le processus de comptage d'une génération  $k \in \mathbb{N}$  de temps  $\{t_i^{(k)} : i \geq 1\}$  est noté

$$N_t^{(k)} = \sum_{i \geq 1} \mathbf{1}_{\{t_i^{(k)} \leq t\}}. \quad (5)$$

Le principe de superposition est le suivant.

1. Pour la première génération  $k = 0$ , on simule des temps

$$\Pi_{(0)} = \{t_1^{(0)}, \dots, t_{N_T^{(0)}}^{(0)}\} \sim \text{PP}(\lambda_0),$$

(processus de Poisson d'intensité  $\lambda_0$ ).

2. Puis, on procède par récurrence sur  $k \geq 1$ . On génère les temps de la génération  $k$  à partir de ceux de la génération  $k - 1$  :

$$\Pi_{(k)} = \{t_1^{(k)}, \dots, t_{N_T^{(k)}}^{(k)}\} \sim \text{PP}\left(\sum_{i \geq 1} g(\cdot - t_i^{(k-1)})\right),$$

(processus de Poisson d'intensité  $\sum_{i \geq 1} g(\cdot - t_i^{(k-1)})$ ).

3. On s'arrête à la génération  $K$  telle que  $N_T^{(K)} = 0$  (à chaque étape, on ne rajoute des temps qu'après les temps précédents, donc au bout d'un certain nombre d'itérations, la nouvelle génération sort complètement de l'intervalle  $[0, T]$ ).
4. On construit alors le processus superposé :

$$\{t_1, \dots, t_{N_T}\} = \bigcup_{k=0}^{K-1} \Pi_{(k)},$$

$$\text{où } N_t = \sum_{k=0}^{K-1} N_t^{(k)}.$$

## ● Extension

Estimer les paramètres du processus Hawkes en calculant l'estimateur au maximum de vraisemblance dans le modèle considéré.

## ● Références bibliographiques

Modèle d'auto-excitation et ses propriétés : [\[Haw71\]](#) [\[Oga98\]](#) [\[Bar63\]](#)

Applications et modélisation : [\[Oga98\]](#) [\[SJ12\]](#)

Simulation : [\[Oga81\]](#)

## 6 Ruines en assurance

### ● Contexte

L'assurance est un contrat de transfert de risques entre deux parties : l'assureur prend à son compte un risque (sinistres, dommages à la personne...) en échange d'une cotisation versée par l'assuré. Les sociétés d'assurance s'appuient sur un grand nombre d'assurés pour mutualiser les risques (effet loi des grands nombres) ; en 2013<sup>13</sup>, le chiffre d'affaires de l'assurance en France dépasse les 180 milliards d'euros. L'évaluation de la solvabilité d'une société d'assurance est une question cruciale et complexe ; en 2008, le groupe américain AIG a dû être renfloué par la Réserve Fédérale Américaine alors qu'il était au bord de la ruine, son capital ayant été largement déprécié suite à la crise des *sub-primes*. En Europe, la réforme réglementaire Solvabilité II<sup>14</sup> impose de mieux mesurer les expositions aux risques, en évaluant par exemple les probabilités d'insolvabilité et de ruine. Afficher des excellents indicateurs de solvabilité sert aussi d'argument marketing auprès des clients.

Plus précisément, une compagnie d'assurance dispose d'un capital (fonds propres) ; elle perçoit régulièrement des cotisations de ses clients, et les indemnise en cas de sinistres. Cette compagnie propose en général des contrats de différents types liés aux différentes branches de son activité : assurance vie, complémentaire santé, assurance habitation, assurance automobile, assurance responsabilité civile, etc., avec des taux de cotisation qui dépendent de la *fréquence* des sinistres et de leur *sévérité*. Elle est ruinée lorsque les montants à rembourser sont supérieurs au capital de départ cumulé avec les cotisations perçues ; cela se produit en cas d'apparition de sinistres plus fréquents ou plus importants (catastrophes naturelles, épidémies...) ou si le capital initial et les cotisations sont insuffisants. Il est donc crucial de déterminer et/ou d'ajuster ces paramètres pour garantir la pérennité d'une compagnie d'assurance. Enfin, il est possible qu'une branche de l'assureur soit au bord de la ruine sans que l'assureur le soit globalement.



### ● Modélisation

Pour commencer, on considère une seule branche d'activité. La compagnie d'assurance a un capital de départ  $x$ , et on note  $R_t$  sa réserve à l'instant  $t$ . La compagnie d'assurance :

---

13. [http://www.ffsa.fr/sites/jcms/p1\\_983486/1-assurance-francaise-demontre-sa-solidite?cc=fn\\_7345](http://www.ffsa.fr/sites/jcms/p1_983486/1-assurance-francaise-demontre-sa-solidite?cc=fn_7345)

14. [https://fr.wikipedia.org/wiki/Solvabilité\\_II](https://fr.wikipedia.org/wiki/Solvabilité_II)

- perçoit des cotisations de ses clients, pour simplifier en continu et uniformément réparties sur l’année. On note  $p$  le taux de cotisation par unité de temps. Les recettes de la compagnie pendant un temps  $t$  sont donc  $pt$ .
- verse des indemnités à ses assurés sinistrés en fonction des dommages subis. On note  $0 = T_0 < T_1 < T_2 < \dots$  les instants aléatoires des sinistres et  $(Z_k)_{k \geq 1}$  les dommages. En posant  $N_t \stackrel{\text{def}}{=} \sum_{k \geq 1} \mathbf{1}_{T_k \leq t}$  le nombre de sinistres entre l’instant 0 et  $t$ , la réserve à la date  $t$  est

$$R_t = x + pt - \sum_{k=1}^{N_t} Z_k.$$

▷ **Modèle, paramètres.** On modélise les montants remboursés  $(Z_k)_{k \geq 1}$  par des lois Gamma i.i.d. de paramètres  $(\alpha, \beta)$ ,  $\alpha > 0$  et  $\beta > 0$ . Les durées entre deux sinistres successifs  $(\tau_k)_{k \geq 1}$  définies par  $\tau_k \stackrel{\text{def}}{=} T_k - T_{k-1}$ , sont supposées indépendantes entre elles et indépendantes des sinistres ; on suppose qu’elles sont distribuées suivant des variables exponentielles de paramètre  $\lambda$ ,  $\lambda > 0$ .

▷ **Ruine.** La société est ruinée à l’horizon  $T$  si la réserve  $R_T$  est négative.

## ● Risque de ruine à l’horizon 1 an

L’objectif de cette partie est de calculer la probabilité de ruine  $\mathbb{P}(R_T \leq 0)$  pour différentes valeurs des paramètres du modèle  $(x, \alpha, \beta, \lambda, p)$  et pour  $T = 1$  an.

1. Estimer la probabilité de ruine au bout d’un an pour les paramètres suivants  $p = 1$ ,  $\alpha = 2.5$ ,  $\lambda = 0.005$ ,  $\beta = 0.12$  et  $x = 100$ . Proposer différentes méthodes de calcul.
2. Déterminer le capital initial que la compagnie doit avoir pour que sa probabilité de ruine à l’horizon annuel soit inférieure à  $10^{-6}$ .
3. Estimer la distribution de  $R_T$  sachant qu’il y a eu ruine à l’instant  $T$ .

## ● Extensions

On traitera l’une des extensions suivantes.

▷ **Risque de ruine à horizon infini.** On veut quantifier si la compagnie d’assurance sera un jour ruinée. Si la société d’assurance est effectivement ruinée, la ruine interviendra à l’instant d’arrivée d’un sinistre : en conséquence, on s’intéresse au processus à temps discret de la somme déboursée échantillonnée aux instants d’arrivée des sinistres. La somme déboursée à l’instant d’arrivée du  $n$ -ième sinistre est égale à  $X_1 + \dots + X_n$ , où  $X_k \stackrel{\text{def}}{=} Z_k - p\tau_k$ . La

probabilité de ruine à horizon infini est donc donnée par

$$\mathbb{P}(\sigma_x < \infty), \quad \text{où} \quad \sigma_x \stackrel{\text{def}}{=} \inf\{n \geq 1, \sum_{k=1}^n X_k > x\},$$

avec la convention  $\inf \emptyset = +\infty$ .

La difficulté dans la simulation en temps infini provient du fait qu'il faut simuler toute la trajectoire  $(\sum_{k=1}^n X_k, n = 0, \dots, +\infty)$  pour déterminer si la réalisation est telle que  $\sigma_x < +\infty$  ou  $\sigma_x = +\infty$ . Pour pallier à cette difficulté, on propose d'effectuer un changement de probabilité exponentiel décrit dans [Gla03] et [BLG09] pour lequel sous une nouvelle probabilité  $\mathbb{P}_\theta$ ,  $\mathbb{P}_\theta(\sigma_x < \infty) = 1$ .

**Quelques résultats théoriques.** On note  $\psi_X$  la fonction génératrice des cumulants de  $X$ . Le livre [Asm03] et l'article [BLG09] rappellent des nombreuses propriétés de la fonction génératrice des cumulants.

Si  $\mathbb{E}[X] < 0$ , il existe une unique valeur positive  $\theta^*$  telle que  $\psi_X(\theta^*) = 0$  : on a alors

$$\mathbb{P}(\sigma_x < +\infty) = \mathbb{E}_{\theta^*} \left[ e^{-\theta^* \sum_{k=1}^{\sigma_x} X_k} \right],$$

où  $\mathbb{E}_\theta$  désigne l'espérance associée à la probabilité  $\mathbb{P}_\theta$ . Voir [Gla03].

**Questions.** On cherchera à

1. estimer la probabilité de ruine pour  $\alpha = 1$ ,  $\beta = 1$ ,  $\lambda = 0.5$ ,  $p = 1$  et différentes valeurs de  $x$  avec la méthode présentée ci-dessus.
2. démontrer certains des résultats énoncés dans cette section.

▷ **Compagnie diversifiée.** Dans cette section, on suppose que la compagnie étudiée propose des assurances habitations et des assurances automobiles, et qu'une partie des sinistres liés aux risques naturels touchent à la fois le secteur automobile et habitation. On ajoute des sinistres propres aux habitations  $N_t^H$  et des sinistres propres aux voitures  $N_t^A$  : le nombre de sinistres affectant la branche habitation jusqu'au temps  $t$  est donc  $N_t^{RN} + N_t^H$ , et le nombre des sinistres affectant la branche automobile est  $N_t^{RN} + N_t^A$ .

1. Estimer le risque de ruine au bout d'un an (on pourra notamment étudier la ruine de chaque branche et les effets de compensation entre branche).
2. Déterminer le capital initial dont doit disposer la compagnie pour que sa probabilité de ruine à horizon 1 an soit inférieure à  $10^{-6}$ . Comment répartir un capital donné entre ses deux branches ?
3. Représenter la distribution de la réserve sachant qu'il y a eu ruine.

## ● Références bibliographiques

Théorie de la ruine : [Asm03].

Changement de probabilité exponentiel : [Gla03][BLG09].



## 7 Barrages et risques d'inondation

Le risque d'inondation est classé premier risque naturel en France par le Ministère de l'Ecologie et du Développement durable<sup>15</sup>. Il concerne 13.300 communes sur le territoire français, dont 300 grandes agglomérations, pour un total de 5 à 6 millions de personnes concernées, directement ou indirectement<sup>16</sup>. Dans le mécanisme de contrôle des risques d'inondation, un rôle est joué par les bassins artificiels, qui ont une fonction de régulation des crues (en plus de la production d'énergie hydroélectrique). En même temps, lorsqu'un bassin artificiel est alimenté par un volume d'eau trop important, pour éviter que le barrage qui le retient ne cède, on est obligé d'ouvrir les vannes d'évacuation de façon importante, ce qui entraîne un fort débit d'eau dans la vallée sous-jacente et donc un risque d'inondation est toujours présent. Il existe en France plusieurs lacs artificiels de grand taille, le lac de Serre-Ponçon en tête, avec ses  $1.2 \times 10^9 m^3$  d'eau retenue.



### • Modélisation



Vidange du barrage, ouverture des vannes.

Dans une région montagneuse, deux vallées sont occupées par deux lacs artificiels créés par deux barrages  $B^1$  et  $B^2$ , dont on modélise le volume contenu  $X_t^1$  et  $X_t^2$  au cours du temps. Chaque lac est alimenté par une rivière principale et plusieurs torrents, qui apportent un volume cumulé  $A_t^i$  ( $i = 1, 2$ ) au temps  $t$ . En même temps, chaque bassin se vide à un taux instantané  $r_i(X_t^i)$  qui est une fonction du volume d'eau  $X_t^i$  stocké dans le bassin (en cas de crue, le lac de Serre-Ponçon peut évacuer jusqu'à  $1.24 \times 10^5 m^3$  par heure). L'équation qui gouverne l'évolution temporelle du contenu de chaque réservoir (unité= $10^5 m^3$ ) s'écrit donc

$$X_t^i = x_0^i + A_t^i - \int_0^t r_i(X_s^i) ds, \quad (6)$$

où  $x_0^i$  est le contenu à l'instant initial. On fait l'hypothèse de débit proportionnel en le volume du barrage :  $r_i(x) = r_i x$ .

15. [http://catalogue.prim.net/43\\_\\_dppr-livretrisqmajeurs-v7.pdf](http://catalogue.prim.net/43__dppr-livretrisqmajeurs-v7.pdf)

16. voir la conférence [http://www.coriolis.polytechnique.fr/Confs/Leleu\\_poster.pdf](http://www.coriolis.polytechnique.fr/Confs/Leleu_poster.pdf)

▷ **Volumes surcritiques.** Chaque barrage est construit pour supporter un volume de remplissage maximal  $x_{\text{crit}}$  (la géométrie de chaque bassin artificiel étant fixée, le niveau critique pourrait aussi bien être exprimé par la hauteur maximale de l'eau à contact du barrage). Au-delà de ce niveau critique, on est obligé d'ouvrir les évacuateurs du barrage à tel point que cela entraîne un risque d'inondation pour la vallée sous-jacente. Pour réduire le plus possible ce risque, on définit le niveau critique  $x_{\text{crit}}^i$  de chaque bassin par le quantile  $Q_{X_T^i}(\alpha)$  d'ordre  $\alpha = 10^{-6}$  à un horizon  $T$ . On souhaite estimer à l'aide de simulations la valeur de ce niveau critique, et évaluer ensuite s'il est envisageable, du point de vue financier et technique, de construire un barrage qui puisse tolérer un tel niveau de remplissage. On cherche aussi à quantifier les risques d'inondation dus à un des deux barrages ou aux deux.

▷ **Modélisation de l'apport d'eau.** On suppose que l'apport d'eau pour chacun des deux bassins est essentiellement déterminé par des chutes de pluie intenses et de courte durée (on néglige les ruissellements continus). Les modèles que l'on peut utiliser pour les précipitations atmosphériques sont typiquement fondés sur l'usage de processus stochastiques à sauts : des travaux de recherche [CIO07] proposent des modélisations détaillées, où l'on décrit l'arrivée d'un épisode de pluie par un processus de Poisson (le début de l'événement correspondant à un instant de saut du processus), la durée de l'épisode par une variable aléatoire positive, et on décrit aussi une structure fine de l'événement entre son début et sa fin. On pourra retenir ici la modélisation aléatoire de l'arrivée des pluies : on fait donc l'hypothèse que le débit entrant dans chaque bassin est bien modélisé par un processus de Poisson composé :

$$A_t^i = \sum_{n=1}^{N_t^i} U_n^i, \quad i = 1, 2 \quad (7)$$

où  $(N_t^1)_t$  et  $(N_t^2)_t$  sont deux processus de comptage d'intensités  $\lambda_1$  et  $\lambda_2$ , et  $(U_n^1)_n$  et  $(U_n^2)_n$  sont des variables aléatoires indépendantes et identiquement distribuées, indépendantes entre elles et indépendantes des processus  $(N^i)_t$ . Pour la loi commune des volumes de pluie  $U_1^1 \sim U_1^2$ , une combinaison de lois exponentielles permettra de modéliser séparément des pluies de grande et petite intensité :

$$\nu(x) = b\delta_1 e^{-\delta_1 x} \mathbf{1}_{\{x>0\}} + (1-b)\delta_2 e^{-\delta_2 x} \mathbf{1}_{\{x>0\}}, 0 < b < 1.$$

## ● Objectifs.

▷ **Etude d'un seul barrage.** Pour commencer, on étudie les barrages séparément. Les questions à aborder sont les suivantes.

- Q1. (a) Estimer le niveau critique  $x_{\text{crit}}^i$  pour chaque bassin, à un horizon  $T=1$  an. Déterminer la distribution du volume d'eau dans chaque bassin lorsqu'il y a un dépassement du niveau critique en  $T$ .
- (b) Pour comprendre si l'évaluation des risques que l'on vient d'effectuer à un horizon fixé est pertinente, on reconsidère le critère de définition du niveau de remplissage critique, en se basant maintenant sur le volume maximum  $\max_{t_j} X_{t_j}$  atteint par le bassin tout au long de l'année. Estimer par simulation le nouveau volume critique, ainsi que la distribution de dépassement de ce niveau au cours de l'année. Comparer avec les résultats précédents.

▷ **Prise en compte des deux barrages, inondation en aval.** Les chaînes de montagnes en amont des deux bassins étant suffisamment distantes, on fera - pour l'instant - l'hypothèse que les précipitations qui alimentent les deux lacs sont indépendantes.

Les deux vallées occupées par les bassins, raides et inhabitées, se rejoignent à l'embouchure d'une vallée plus large qui est occupée par des villages. Au point de jonction  $O$  entre les deux vallées, une rivière est formée par l'eau venant des deux bassins, qui atteint le point  $O$  après avoir parcouru la première vallée en un temps  $T_1$ , et la deuxième en un temps  $T_2$ . On supposera que ces retards d'arrivée d'eau en  $O$  sont fixes, et indépendants du volume d'eau qui se déplace. Au point de jonction des vallées, le débit observé au temps  $t$  est donc

$$D_t = r_1(X_{t-T_1}^1) + r_2(X_{t-T_2}^2).$$

La rivière qui coule dans la vallée et qui est alimentée par les deux barrages, peut tolérer un débit limite, au-delà duquel la rivière déborde et entraîne des inondations des zones habitables. On traitera les questions suivantes.

- Q2. (a) Quelle est la probabilité de dépasser le débit de seuil dans la vallée, soit  $\mathbb{P}(D_T \geq \text{seuil})$  ?
- (b) Quelle est la probabilité de dépassement du débit de seuil lorsque l'un des deux bassins est rempli au-dessus du niveau critique ?
- (c) A partir de l'observation du débit instantané au point de jonction  $O$ , on souhaite remonter à de l'information sur l'état de remplissage des bassins. Estimer le volume d'eau dans chaque bassin lorsqu'on observe un niveau limite de débit à la jonction.

## ● Extensions

On cherchera à aborder l'un des points suivants.

- **Dépendance** : On souhaite relaxer l'hypothèse d'indépendance des précipitations en amont des deux bassins (i.e. l'indépendance des deux

processus de comptage des pluies). Proposer une modélisation qui tienne compte d'une dépendance entre les phénomènes atmosphériques. On reprendra l'analyse précédente dans ce cadre.

- **D'autres sources d'aléa** : La vallée habitée est alimentée en eau pas seulement par les deux bassins, mais aussi par l'arrivée de pluies. Prendre en compte ce phénomène dans la modélisation et l'analyse du risque d'inondation.
- **Une analyse en temps long** : Proposer une analyse du système (6)-(7) en temps long ( $t \rightarrow \infty$ ), en étudiant notamment la possibilité d'obtenir des formules explicites pour la loi de  $X_t^i$  dans ce régime, et donc pour ses quantiles. Comparer avec les simulations de la Q1.

## ● Références bibliographiques

Modèles de barrage : [Asm03].

Modèles de pluie : [Sam04][CIO07].

## 8 Equation de la chaleur et résolution Monte-Carlo multi-niveaux (pour binôme mixte SNA-SNE)

### • Problème et modélisation

De nombreux phénomènes physiques sont décrits par un modèle mathématique, constitué souvent d'équations aux dérivées partielles (EDP), qui de façon remarquable ont une représentation probabiliste. Par exemple en *thermique*, la répartition stationnaire de la température  $u$  dans un milieu thermiquement conducteur est décrite par la loi de Fourier  $q = -K\nabla u$  reliant le flux de chaleur  $q$  au gradient de température ( $K$  étant la conductivité thermique du milieu). La conservation de l'énergie s'exprime en fonction des sources de chaleur  $c$  par  $\operatorname{div}(q) = c$ , ce qui donne au final une équation de type

$$-\operatorname{div}(K\nabla u) = c.$$

Il convient d'adjoindre à cette équation des conditions aux limites du milieu  $D \subset \mathbb{R}^d$  (la dimension physique est  $d = 3$ ), imposant par exemple que la température au bord  $\partial D$  est égale à  $g$ , soit

$$u = g \text{ sur } \partial D.$$

Nous supposons ici une conductivité thermique constante  $K$  : ainsi les équations aboutissent à l'*équation de chaleur avec condition au bord de Dirichlet*

$$\frac{1}{2}\Delta u = f \text{ dans } D, \quad u = g \text{ sur } \partial D, \quad (8)$$

où  $f := -\frac{c}{2K}$  et où l'écriture avec le facteur  $\frac{1}{2}$  est choisie par convention en vue de la représentation probabiliste qui suit.

Depuis les années 50 et notamment les travaux de Kac [Kac49], on sait représenter la solution  $(u(x); x \in D)$  comme l'espérance d'une fonctionnelle de la trajectoire du mouvement brownien <sup>17</sup> $W = (W_t)_{t \geq 0}$  partant de  $x \in D$  en  $t = 0$  et arrêté au premier instant  $\tau$  où il touche la frontière  $\partial D$ . Cela s'écrit

$$u(x) = \mathbb{E}_x \left( g(W_\tau) - \int_0^\tau f(W_s) ds \right). \quad (9)$$

Cette représentation, qui porte le nom de formule de Feynman-Kac, permet alors d'évaluer ponctuellement  $u(x)$  à l'aide d'un grand nombre de simulations Monte-Carlo des trajectoires browniennes.

---

17. processus  $W$  continu  $d$ -dimensionnel, à composantes indépendantes, dont chaque composante  $W^{(k)}$  est à accroissements indépendants avec  $W_t^{(k)} - W_s^{(k)}$  ( $t > s$ ) de loi gaussienne centrée de variance  $t - s$ .

## ● Objectifs

Il s'agit de tester un algorithme Monte-Carlo qui calcule  $u(x)$  en tout point  $x$ , à l'aide de la méthode multi-niveaux exposée dans [Hei01] : cette approche combine un procédé d'approximation de fonction sur des grilles de points de  $D$  et des simulations de trajectoires partant de ces points, le tout étant couplé via un raffinement progressif<sup>18</sup> des grilles.

Les points à aborder sont les suivants :

- s'approprier les outils et résultats développés dans [Hei01],
- comparer avec une méthode de discrétisation d'EDP (partie SNE).

A propos du domaine  $D$ , on pourra prendre une géométrie simple (approximativement rectangulaire), inspirée par exemple d'un *casert d'élève* ou d'une *salle de cours*. La source de chaleur décrite par la fonction  $f$  sera par exemple celle d'un radiateur ou d'une fenêtre ouverte vers l'extérieur froid. On pourra commencer les tests en dimension  $d = 2$  avant de poursuivre en dimension 3.

La simulation des temps de sortie du mouvement brownien pourra s'inspirer de discrétisation simple du mouvement brownien ou de procédures plus avancées comme décrites dans [GM10][Gob3A].

## ● Références bibliographiques

Sur les méthodes multi-niveaux : [Hei01].

Sur la simulation du mouvement brownien et des temps de sortie : [GM10][Gob3A].

Sur les formules de Feynman-Kac : [Kac49] [KS91].

---

18. par niveaux

## 9 Solutions approchées certifiées pour l'équation de Burgers visqueuse et incertitudes (pour binôme mixte SNA-SNE)

Le but de ce projet est de reprendre les idées présentées dans [JNP10] et [CGJ<sup>+</sup>] pour implanter une méthode de base réduite certifiée pour l'équation de Burgers visqueuse.

L'ambition de ce projet est de concevoir une méthode numérique qui permette l'analyse de sensibilité de paramètres connus de façon imprécise (viscosité, conditions initiales, etc.) dans une équation de la mécanique des fluides. Afin de simplifier le problème, on considèrera l'équation de Burgers visqueuse 1D (au lieu des équations de Navier-Stokes 3D!). Après lecture approfondie de [JNP10] on implantera la méthode de type éléments finis pour résoudre l'équation de façon précise et l'on en extraira les bases réduites qui permettront une résolution accélérée indispensable pour l'exploration de la solution lorsque l'on fait varier différents paramètres (par exemple la viscosité ou les conditions au bord ou initiales...).

Enfin, on assignera des lois de probabilité aux paramètres incertains et on construira des intervalles de confiance sur la fonction solution de l'EDP, à l'aide des outils mis en place avant. Ce procédé permet de donner des *indices de confiance* sur les quantités liées à la solution cherchée (comme c'est couramment fait par exemple lors de prévisions météorologiques).

En parallèle, on tâchera de voir si les méthodes issues de [CGJ<sup>+</sup>] pour l'analyse de sensibilité peuvent s'appliquer dans ce contexte.

### ● Références bibliographiques

Pour les bases certifiées : voir [JNP10].

Pour les sensibilités : voir [CGJ<sup>+</sup>].

## Références

- [Asm03] S. Asmussen. *Applied probability and queues*, volume 51 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2003. Stochastic Modelling and Applied Probability.
- [Bar63] M.S. Bartlett. The spectral analysis of point processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 264–296, 1963.
- [BEGK04] A. Bovier, M. Eckhoff, V. Gaynard, and M. Klein. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4) :399–424, 2004.
- [BLG09] J.H. Blanchet, K. Leder, and P.W. Glynn. Efficient simulation of light-tailed sums : an old-folk song sung to a faster new tune. . . In *Monte Carlo and quasi-Monte Carlo methods 2008*, pages 227–248. Springer, Berlin, 2009.
- [Bou02] P. Bougerol. *Processus de sauts et files d’attente*. Cours de M1 en téléchargement libre à <http://www.proba.jussieu.fr/supports.php>, 2002.
- [BR78] J. Bussac and P. Reuss. *Traité de neutronique*. Hermann, Paris, 1978.
- [Buc04] J. Bucklew. *Introduction to Rare Event Simulation*. Springer, 2004.
- [CDL<sup>+</sup>89] M. Cessenat, R. Dautray, G. Ledanois, P.L. Lions, E. Pardoux, and R. Sentis. *Méthodes probabilistes pour les équations de la physique*. Collection CEA, Eyrolles, 1989.
- [CGJ<sup>+</sup>] T. Cartailier, A. Ghaus, A. Janon, H. Monod, C. Prieur, and N. Saint-Geours. Sensitivity analysis and uncertainty quantification for environmental models.
- [CIO07] P. Cowpertwait, V. Isham, and C. Onof. Point process models of rainfall : developments for fine-scale structure. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 463(2086) :2569–2587, 2007.
- [FS02] H. Föllmer and A. Schied. *Stochastic finance*, volume 27 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2002. An introduction in discrete time.
- [GHS00] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Variance reduction techniques for estimating value-at-risk. *Management Science*, 46 :1349–1364, 2000.
- [Gla03] P. Glasserman. *Monte Carlo methods in Financial Engineering*. Springer Verlag, New York, 2003.



- [GM10] E. Gobet and S. Menozzi. Stopped diffusion processes : boundary corrections and overshoot. *Stochastic Processes and Their Applications*, 120 :130–162, 2010.
- [Gob3A] E. Gobet. *Méthodes de Monte-Carlo et processus stochastiques : du linéaire au non-linéaire*. Cours de l’Ecole Polytechnique, MAP564, 3A.
- [Haw71] Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1) :83–90, 1971.
- [Hei01] S. Heinrich. Multilevel Monte Carlo Methods. In *LSSC ’01 Proceedings of the Third International Conference on Large-Scale Scientific Computing*, volume 2179 of *Lecture Notes in Computer Science*, pages 58–67. Springer-Verlag, 2001.
- [HTB90] P. Hanggi, P. Talkner, and M. Borkovec. Reaction-rate theory : fifty years after Kramers. *Rev. Modern Phys.*, 62(2) :251–341, 1990.
- [JNP10] A. Janon, M. Nodet, and C. Prieur. Certified reduced-basis solutions of viscous Burgers equation parametrized by initial and boundary values. *ArXiv e-prints*, October 2010.
- [Kac49] M. Kac. On distributions of certain Wiener functionals. *Trans. Amer. Math. Soc.*, 65 :1–13, 1949.
- [Kra40] H.A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7, 1940.
- [KS91] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*. Springer Verlag, second edition, 1991.
- [LPS98] B. Lapeyre, E. Pardoux, and R. Sentis. *Methodes de Monte Carlo pour les processus de transport et de diffusion*. Collection Mathématiques et Applications 29 - Springer Verlag, 1998.
- [Min11] A. Minca. *Modélisation mathématique de la contagion de défaut*. PhD thesis, Université Pierre et Marie Curie, 2011.
- [Mou11] A. Moussa. *Contagion and Systemic Risk in Financial Networks*. PhD thesis, Columbia University, 2011.
- [Oga81] Y. Ogata. On lewis’ simulation method for point processes. *Information Theory, IEEE Transactions on*, 27(1) :23–31, 1981.
- [Oga98] Y. Ogata. Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2) :379–402, 1998.
- [OV05] E. Olivieri and M.E. Vares. *Large deviations and metastability*, volume 100 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005.

- [Pap75] G. Papanicolaou. Asymptotic analysis of transport processes. *Bulletin of the American Mathematical Society*, 81(2) :330–392, 1975.
- [Pot] R. Poteau. *Modélisation moléculaire*. Cours de M1 en téléchargement libre à [http://www.ressources-pedagogiques.ups-tlse.fr/cpm/MODMOL/slide\\_mod\\_mol-C.pdf](http://www.ressources-pedagogiques.ups-tlse.fr/cpm/MODMOL/slide_mod_mol-C.pdf).
- [PW89] S. Parekh and J. Walrand. A quick simulation method for excessive backlogs in networks of queues. *IEEE Trans. Automat. Control*, <http://www.ece.ucdavis.edu/~chuah/classes/EEC273/refs/PW89-quick-simulation.pdf>, 34 :54–66, 1989.
- [Sam04] S. Sambou. Modèle statistique des hauteurs de pluies journalières en zone sahélienne : exemple du bassin amont du fleuve sénégal. *Hydrological Sciences – Journal des Sciences Hydrologiques*, 49(1) :115–129, 2004.
- [SJ96] G.D. Smith and R.L. Jaffe. Quantum chemistry study of conformational energies and rotational energy barriers in n-alkanes. *The Journal of Physical Chemistry*, 100(48) :18718–18724, 1996.
- [SJ12] A. Simma and M.I. Jordan. Modeling events with cascades of Poisson processes. *arXiv preprint arXiv :1203.3516*, 2012.
- [SW95] A. Schwartz and A. Weiss, editors. *Large Deviations for Performance Analysis*. Chapman and Hall, London, 1995.
- [Wik] Wikipedia. M/M/1 queue. [http://en.wikipedia.org/wiki/M/M/1\\_queue](http://en.wikipedia.org/wiki/M/M/1_queue).